1645

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

JUL 09 2004

TECH CENTER 1600/2900

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/880,748	06/15/2001	Steven M. Ruben	PF523P1	5654

22195 7590 06/16/2004

HUMAN GENOME SCIENCES INC
INTELLECTUAL PROPERTY DEPT.
14200 SHADY GROVE ROAD
ROCKVILLE, MD 20850

EXAMINER	
DUFFY, PATRICIA ANN	
ART UNIT	PAPER NUMBER
1645	

DATE MAILED: 06/16/2004

Please find below and/or attached an Office communication concerning this application or proceeding.

PTO-90C (Rev. 10/03)

Office Action Summary

Application No.	Applicant(s)	
09/667,130	BARNWELL, JOHN W.	
Examiner	**Art Unit**	
Patricia A. Duffy	1645	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *26 September 2003*.

2a)☒ This action is **FINAL**. 2b)☐ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *22 and 23* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *22 and 23* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☒ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08) Paper No(s)/Mail Date *attachment*.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____.
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☐ Other: _____.

RESPONSE TO AMENDMENT

The response filed September 26, 2003 has been entered into the record. Claims 22 and 23 are pending and under examination, all other claims having been canceled.

The text of Title 35 of the U.S. Code not reiterated herein can be found in the previous office action.

Rejections Maintained

The preliminary amendment filed 9-21-00 stands objected to under 35 U.S.C. 132 because it introduces new matter into the disclosure. 35 U.S.C. 132 states that no amendment shall introduce new matter into the disclosure of the invention. The added material which is not supported by the original disclosure is as follows: " EcoRI digest of purified ..... for 2 months." which is inserted at page 5, line 18 after "1989". The objection is maintained for reasons made of record in Paper No. 8, mailed 7-1-02 and the last office action of record.

Applicant's arguments have again been considered but are again not persuasive. Applicant argues Dr. Crothers declaration that opines that the conditions specifically disclosed in the Southern et al article would be considered "stringent". This is not persuasive again, the incorporation by reference must be specific. Declarant is basically attesting that any condition of hybridization is stringent. This is not persuasive, Southern et al does not define these conditions as "Stringent". Further, Southern et al is a guidebook to selection of general hybridization conditions and does not define any specific conditions as stringent. The specification does not direct one to any specific conditions set forth in Southern et al and erroneously indicates that Southern et al defines stringent hybridization conditions, it does not. The incorporation of the specific material inserted into the specification is not a result of a specific incorporation by reference by pointing in

particular to the relied upon teachings in Southern et al, but instead reflects an improper general incorporation by reference to any condition recited therein. Nowhere in Southern et al is the word "Stringent" used. Therefore, nothing in Southern et al teaches "stringent conditions" as specifically referenced by the specification. As such, the skilled artisan would not be pointed to any particular hybridization conditions recited by Southern et al, none are characterized by Southern et al as "stringent" as opposed to not stringent. As previously indicated, Southern et al is merely a guidebook for selection of conditions that allow for hybridization between two nucleic acids and these conditions are dependent on the particular structure (G+C) content of the nucleic acid under study. As such, there is nothing articulated in the specification to any particular set of conditions in Southern et al. The declaration of Dr. Crothers is not persuasive on this point, *it is the specification that must particularly point to those conditions which are considered incorporated by reference.* This point is made completely clear in the office actions of record. Applicant argues that Dr. Crothers' conclusion with regard to experimentation conditions is based upon his examination of the experimental conditions explored and used in the Southern et al publication. This again is not persuasive, Dr. Crothers' opinion is not supported by extrinsic evidence for the recited nucleic acid and moreover supports the assertion of the examiner that the incorporation by reference is not specific to any subject material in Southern et al. This argument appears to support that the reference is an improper general incorporation and does not in fact specifically reference the incorporated material as required by the MPEP and Hawkins decisions as previously set forth in the all the office actions of record. Unlike *In re Voss*, the incorporation is not reasonably precise, it was not made for clarity and appears erroneous on its face. It specifically references the Southern et al "as defined by", but the reference did not define such conditions, but recites multiple conditions. This application is unlike the situation in *In re Voss*, because in Voss, Applicant pointed to a the relied upon material "discussion of glass-ceramic material and their production". The incorporation by reference in this specification does not. The

entirety of Southern et al is drawn to selection of hybridization conditions and the variation of those to achieve specific hybridization. There is no discussion of stringency and no particular conditions referenced that are particularly defined as such. It is the specification that must clearly and unambiguously point to the incorporating material. Unlike the relied upon paragraph in *In re Voss*, that does such, this specification does not for all the previous reasons made of record. Applicant argues that stringent conditions are not empirical in nature and are in fact a combination of temperature and salt conditions. Applicant argues that the concept of stringency (i.e. the intrinsic specificity of the hybridization reaction depends on the annealing conditions employed was familiar to those of ordinary skill in the field. While this may be true, the concept of "stringency" does not define any particular conditions, it is *relative*. Kennellet al "Principles and properties of nucleic acid hybridization", Progr. Nucl. Acid Res. Mol. Biol. 11:259-301, 1971) teaches that ". It should be emphasized that the extent to which mismatched hybrids can form is relative and dependent upon the "stringency" of the reaction conditions. The stringency refers to the extent to which the reaction conditions allow only completely complementary structures to form. Generally, stringency is proportional to temperature and inversely proportional to salt conditions." (page 293, first full paragraph). Thus, stringent conditions are not a particular set of defined conditions because stringency is a relative concept and "stringent conditions" can vary depending upon the temperature and salt conditions. The skilled artisan would not pick one set of conditions over any other set in view of the teachings of Southern et al because the specification does not teach nor does it contemplate the degree of mismatch contemplated. As such, this specification does not point to any particular conditions that it defines or Southern et al defines as "stringent". In contrast to Applicants assertions, the Southern et al reference uses multiple conditions of temperature in salt (see figures) and as such does not define any particular conditions. There are no one set of conditions that are "defined as stringent conditions". Even if one were to consider all the conditions set forth in Southern et al. as

"stringent", then Applicant's incorporation by reference falls to the level of a improper general incorporation by reference. Applicant argues that as evidence that "stringent hybridization conditions" are not empirical in nature and need not be independently determined for every different nucleic acid, Applicants point to the Written description guidelines Example 9. This is not persuasive, it is relying upon hypothetical teachings. All of the Examples are hypothetical in nature. Declarant and Kennel's asserted concept of stringency emphasizes the relative nature of "stringency" and that no particular conditions are art defined as stringent, nor does Southern et al define such. As to the nature of empirical nature of hybridization, the property of G+C content on Tm etc are all exemplified by Kennel as cited *supra*. The structure of the nucleic acids completely reflects its ability to form a complex with a target under specifically defined conditions. However, this specification does not reference with any particularity any specifically defined conditions. It relies upon an vague incorporation by reference utilizing language not present in the reference and the skilled artisan is left guess as to the meaning thereof and the particular conditions referenced. Declarant's opinion is not persuasive on this issue because it essentially states that all hybridization conditions are stringent, while the art defines stringency as a relative concept and not defined by any particular art defined conditions.

The rejection is maintained for reasons made of record herein and all the previous reasons made of record.

Claims 22 and 23 are rejected under the judicially created doctrine of obviousness-type double patenting as being unpatentable over claims 22, 26 and 28 of copending Application No. 08/719,821, now US Patent No. 6,706,872. Although the conflicting claims are not identical, they are not patentably distinct from each other because the species of full complements claimed in the copending application, would by definition hybridize to SEQ ID NO:1 and thus anticipate the instant genus claims . The

rejection is maintained for reasons made of record in Paper No. 8, mailed 7-1-02. Applicants indicate that a terminal disclaimer would be filed upon indication of allowable subject matter. The rejection is hereby maintained until resolution by terminal disclaimer.

Claim 23 stands rejected under 35 U.S.C. 112, first paragraph, as containing subject matter which was not described in the specification in such a way as to reasonably convey to one skilled in the relevant art that the inventor(s), at the time the application was filed, had possession of the claimed invention. The rejection is maintained for reasons made of record in Paper No. 8, mailed 7-1-02.

Applicants arguments have been carefully considered but are not persuasive. Applicants argue that the general incorporation by reference of Southern et al that recites numerous hybridization conditions and in view of the Declaration of Dr. Crothers is not persuasive. Applicant's arguments with respect to the insertion of new matter into the specification were not persuasive for reasons set forth above.

Claim 22 stands rejected under 35 U.S.C. 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention. The rejection is maintained for reasons made of record in Paper No. 8, mailed 7-1-02.

Applicants argue that this is not well taken and should be withdrawn. This is not persuasive for reasons made of record previously and in view of the concept of stringency is relative and that Southern et al does not define any condition that is "stringent". This is not persuasive for reasons made of record previously and Applicant's arguments with respect to the insertion of new matter into the specification were not persuasive for reasons set forth above.

Claims 22 and 23 stand rejected under 35 U.S.C. 102(b) as being anticipated by Sigma Molecular biology Product Guide, 1991, pages 54-56). The rejection is maintained for reasons made of record in Paper No. 8, mailed 7-1-02.

Applicants arguments have been carefully considered but are not persuasive. Applicants argue that there is no mention of stringent hybridization conditions in Sigma at all and therefore it can not anticipate the claimed invention. This is not persuasive, the claims are drawn to products by function. Applicants also argue that the rejection is incompatible with the 112, first paragraph rejection of record. This is not persuasive because as Applicants well know, the rejections are made under different statutes that have different specific requirements. The products of the art inherently have the recited function of hybridization based upon their 100% identity across the recited nucleotides. Applicants have not provided any extrinsic evidence that the products of the art would not hybridize under any specific set of allegedly "stringent" conditions.

Since the Office does not have the facilities for examining and comparing applicant's product with the product of the prior art, the burden is on applicant to show a novel or unobvious difference between the claimed product and the product of the prior art (i.e., that the nucleic acid of the prior art does not possess the same functional characteristics of the claimed nucleic acid). See In re Best, 562 F.2d 1252, 195 USPQ 430 (CCPA 1977) and In re Fitzgerald et al., 205 USPQ 594.

Claims 22 and 23 are rejected under 35 U.S.C. 112, first paragraph, as containing subject matter which was not described in the specification in such a way as to reasonably convey to one skilled in the relevant art that the inventors, at the time the application was filed, had possession of the claimed invention. The following *written description* rejection is set forth herein.

Applicants arguments have been carefully considered but are not persuasive. Applicants argue a Declaration by the inventor Dr. Barnwell. Dr. Barnwell attests that SEQ ID NO:2 is in fact the full-length open reading frame and attempts to explain the differences between the apparent molecular weights of the protein as compared with the calculated molecular weight. This is not persuasive, the specification specifically teaches that it is a fragment. This specification the teachings of which are sworn and attested to by Dr. Barnwell. Further, Dr. Barnwell is attesting that they have disclosed the full-length open reading frame of 1011 amino acids and that the sequence of depicted in SEQ ID NO:2 is the complete open reading frame. This is not persuasive on its face. SEQ ID NO:2 is 1018 amino acids in length and is described as a partial sequence. Applicant is arguing that the full length is 1011 amino acids long. This argument is entirely incompatible with the written description of the disclosure. Applicants are attempting to redefine the invention as something that is shorter than that described by SEQ ID NO:2 which is 1018 amino acids long. It is the specification that must teach conception of what is the invention. The specification teaches that the genomic fragment is an incomplete reading frame. Declarant, discovering his error in characterization well after filing this application can not seek to "correct" the specific written description thereof of the specification as filed, by assertion that the nucleic acid of SEQ ID NO:1 is the full-length open reading frame, when the specification clearly indicates otherwise. The declaration of Dr. Barnwell, in essence an attempt to redefine the described invention of the specification after the filing date. The specification does not state that the complete gene codes for 1011 amino acids with a calculated molecular weight of 112,725 Da. There is no recognition of this fact in the specification as filed. Further, one skilled in the art would take the characterization of the gene as incomplete in view of the lack of a recited conventional start codon on the protein, the discrepancy of the molecular weights. Declarant provides many different possible explanations, but no extrinsic evidence that the particular system used provides for a protein having these characteristics. Further, the inability of a

indicating that a misrepresentation in a patent specification will only defeat the enablement of an invention if the skilled artisan fails to recognize the mistake and relies on it to practice the invention. This is not persuasive, the question of Wahl involves a best mode inquiry and whether the inventor failed to disclose particular manufacturing procedures beyond the information sufficient for enablement and that such failure gives rise to an inference that he concealed information which one of the ordinary skill in the art would not know. The issue at hand is conception of the full-length open reading frame by way of written description and not concealment of best mode or enablement. Further, the missing descriptive information is not a routine manufacturing choice. For written description of nucleic acids, the nucleic acid itself is required. See Fiers v. Revel, 25 USPQ2d 1601, 1606 (CAFC 1993) and Amgen Inc. V. Chugai Pharmacentical Co. Ltd., 18 USPQ2d 1016. This written description inquiry is independent from enablement Vas-Cath Inc. v. Mahurkar, 19 USPQ2d 1111, makes clear that "applicant must convey with reasonable clarity to those skilled in the art that, as of the filing date sought, he or she was in possession *of the invention*. The invention is, for purposes of the 'written description' inquiry, *whatever is now claimed*." (See page 1117.) The specification conveys that the nucleic acid of SEQ ID NO:1 encoding the protein of SEQ ID NO:2 is not the full-length open reading frame. The specification does not clearly allow persons of ordinary skill in the art to recognize that the inventors invented the full length open reading frame as argued, specifically teaches to the contrary and one skilled in the art would not have any specific framework of reference to believe otherwise, since other similar proteins and nucleic acids have not been disclosed in the art. Therefore, unlike Example 8, there is no frame of reference for either completeness of the sequence or function of the protein encoded by the nucleic acid, and the teachings of the specification would necessarily be taken as true. Applicants argue that because there is a single species and reduction to practice of a single species is representative of the genus because one skilled in the art would not expect substantial variation in the sequence. This is not persuasive, as

bacterium to splice any introns, would apparently provide for a protein of increased rather than decreased molecular weight because the intervening sequence (introns) would be expressed by the bacterium. Declarant provides no evidence that the this is the case in the described expressed by the method set forth in specification. With respect to in situ proteolysis of the expressed protein, this is hypothetical and does not provide extrinsic evidence that the protein produced by the method described in the specification undergoes *in situ* proteolysis. There is not a single post-translational modification described in the specification with respect to this protein. Therefore, the discrepancy of the molecular weights because of the inability of the bacterium to perform certain post-translational modifications is hypothetical. No posttranslational modifications of the protein encoded by the nucleic acid are taught nor contemplated by this specification. Declarant also attests that discrepancies are common among *Plasmodium* proteins because they contain repeated proline rich amino acid motifs that render them less susceptible to denaturation by SDS. This is not persuasive, there is no evidence to support this problem associated with SEQ ID NO:2. Applicants submit that the instant specification enables the complete open reading frame. Enablement is not the issue here. Written description of the full length open reading frame and whether or not, Applicants at the time of filing, as described by the specification at the time of filing, conveyed possession of the full-length open reading frame. The examiner contends that the written description of the specification did not, convey possession of the full-length open reading frame and as such is not analogous to Example 8. Further, in contrast to Example 8, no asserted function is associated with the protein fragment of SEQ ID NO:2 and coding for a surface protein of *P. vivax* is not a function of the protein, it is not a description of what it does, but where it is located. Location is not equivalent to function of a protein. Applicants were not in possession of the genus comprising SEQ ID NO:1 because they do not describe the full-length open reading frame, nor do they describe the function of the protein. Applicants argues *Wahl Instrument v Acvious Inc.* 950 F2.d 1575 21 USPQ 2d 1123 (Fed Cir. 1991)

previously set forth, the claims encompass variants that do not encode proteins. The specification does not disclose that the disclosed species of SEQ ID NO:1 was used in any manner to hybridize to other nucleic acids. Unlike Example 9 of the written description guidelines, there is no disclosed example where SEQ ID NO:1 is used for the isolation of hybridizing nucleic acids. Additionally, the methods of Example 9 used defined conditions and recite a specific functional activity and were demonstrated to have the asserted activity. No such information is provided in this specification, no specific function is disclosed, SEQ ID NO:1 was not used to isolate hybridizing sequences and as such the specie of SEQ ID NO:1 is not representative of a genus of hybridizing sequences with identical biological function. Therefore, the skilled artisan given the written description of the specification would clearly recognize that Applicants had no conception by way of written description of the genus of nucleotides hybridizing under stringent conditions, nor were they in possession, of the full-length open reading frame.

Status of Claims

Claims 22 and 23 stand rejected.

Conclusion

THIS ACTION IS MADE FINAL. Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any

PTO/SB/08a/b (08-03)
Approved for use through 07/31/2006. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

Substitute for form 1449A/B/PTO

INFORMATION DISCLOSURE STATEMENT BY APPLICANT

(Use as many sheets as necessary)

Sheet 1 of 2

Complete if Known	
Application Number	09/667,130-Conf. #8596
Filing Date	September 21, 2000
First Named Inventor	John W. Barnwell
Art Unit	1645
Examiner Name	P. Duffy
Attorney Docket Number	05986/1007686-US5

U.S. PATENT DOCUMENTS

Examiner Initials*	Cite No.[1]	Document Number: Number-Kind Code[2] (if known)	Publication Date MM-DD-YYYY	Name of Patentee or Applicant of Cited Document	Pages, Columns, Lines, Where Relevant Passages or Relevant Figures Appear
PAD	AA**	US-5,001,225-B1	03-19-1991	Taylor	
PAD	AB**	US-5,130,416-B1	07-14-1992	Wellems et al.	

FOREIGN PATENT DOCUMENTS

Examiner Initials*	Cite No.[1]	Foreign Patent Document: Country Code[3]-Number[4]-Kind Code[5] (if known)	Publication Date MM-DD-YYYY	Name of Patentee or Applicant of Cited Document	Pages, Columns, Lines, Where Relevant Passages or Relevant Figures Appear	T[6]

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant. **CITE NO.: These patent(s) or publication(s) which are marked with an double asterisk (**) next to the Cite No. are not supplied because they were previously cited by or submitted to the Office in a prior application relied upon in this application for an earlier filing date under 35 U.S.C. 120. [1] Applicant's unique citation designation number (optional). [2] See Kinds Codes of USPTO Patent Documents at www.uspto.gov or MPEP 901.04. [3] Enter Office that issued the document, by the two-letter code (WIPO Standard ST.3). [4] For Japanese patent documents, the indication of the year of the reign of the Emperor must precede the serial number of the patent document. [5] Kind of document by the appropriate symbols as indicated on the document under WIPO Standard ST.16 if possible. [6] Applicant is to place a check mark here if English language Translation is attached.

NON PATENT LITERATURE DOCUMENTS

Examiner Initials*	Cite No.[1]	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc.), date, page(s), volume-issue number(s), publisher, city and/or country where published.	T[2]
PAD	CA**	Aikawa, M. et al., Am.J. Pathol. 79:285, 1975.	
	CB**	Anderson, R.G.W., PNAS (USA) 90:10909, 1993.	
	CC**	Anderson, R.G.W., et al., Science 255:410, 992.	
	CD**	Atkinson, C.T. et al., Blood Cells 16:351, 1990.	
	CE**	Barnwell, J.W. et al., Experimental Parasitology 70:85, 1990.	
	CF**	Bianco, A.E. et al., PNAS (USA) 83:8713, 1986.	
	CG**	Campbell, A.M., in "Laboratory Techniques in Biochemistry and Molecular Biology", Elsevier, 1991, pp. 20-27.	
	CH**	Chappell, T.G. et al., Cell 45:3, 1986.	
	CI**	Eckert et al., Exp. Parasitol. 75:323, 1992.	
	CJ**	Galinski et al., Cell 69:1213, 1992.	
	CK**	Harlow, E. et al., "Antibodies - A Laboratory Manual", Cold Spring Harbor Laboratory, 1988.	
	CL**	Hellstrom, K.E. et al., in "Monoclonal Antibodies for Cancer Detection and Therapy", edited by R.W. Baldwin, Academic Press, 1975, p. 20.	
	CM**	Howard et al., J. of Cell Biol. 103:1269, 1986.	
	CN**	James, M.A. et al., Abstracts of the 41st Annual Meeting of the American Society of Tropical Medicine and Hygiene, Seattle, Wash. Nov. 16-19, 1992, Abstract No. 135, p. 145.	
	CO**	Kumar, N. et al., PNAS (USA) 85:6277, 1988.	
	CP**	Lathe, R., J. Mol. Biol. 183:1, 1985.	
	CQ**	Matsumoto, Y. et al., Am. J. Trop. Med. Hyg. 39(4):317, 1988.	
	CR**	Maurer, P.H. et al., Methods in Enzymology, 70:49, 1980.	
	CS**	Minchiotti, L. et al., Abstract Biochim, Biophys. Acta, 1119:232, 1992.	
	CT**	Nolte, D., et al., Mol and Biochem. Parasitol. 49:253, 1991.	
	CU**	O'Connell et al., MD & DI pp. 31-36, December 1985.	
	CV**	O'Connell t al., Clin. Chem 31(9):1424, 1985.	

Examiner Signature	PATRICIA A DUFFY	Date Considered	6/8/04

extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the mailing date of this final action.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Patricia A. Duffy whose telephone number is 571-272-0855. The examiner can normally be reached on M-F 6:30 pm - 3:00 pm. If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Lynette Smith can be reached on 571-272-0864.

The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Patricia A. Duffy

Primary Examiner

Art Unit 1645

PTO/SB/08a/b (08-03)
Approved for use through 07/31/2006. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Substitute for form 1449A/B/PTO

INFORMATION DISCLOSURE STATEMENT BY APPLICANT

(Use as many sheets as necessary)

Sheet 2 of 2

Complete if Known	
Application Number	09/667,130-Conf. #8596
Filing Date	September 21, 2000
First Named Inventor	John W. Barnwell
Art Unit	1645
Examiner Name	P. Duffy
Attorney Docket Number	05986/1007686-US5

PAD	CW**	Panton et al., Mol. and Biochem. Parasitol. 35:149, 1989.	
	CX**	Rock et al., Parasitol. 95L209, 1987.	
	CY**	Thomas, A.W. et al. Parasite Immunol. 12:105, 1990.	
	CZ**	Vernick et al., Nucleic Acids Res. 16(14):6883, 1988.	
	CA1**	Wellems et al., Proc. Natl. Acad. Sci. USA 83:6065-6069, 1986.	
	CB1**	Wilson et al., Parasitol. 71:183, 1975.	
	CC1**	Wilson et al., The Lancet, July 26, 1969, pp. 201-204.	
	CD1**	Wilson et al., Nature 284:451, 1980.	
	CE1**	Wilson et al., Int. J. Parasitol. 3:511, 1973.	
	CF1**	Yang Y-F et al., Mol. and Biochem. Parasitol. 26:61, 1987.	

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant. **CITE NO.: Those patent(s) or publication(s) which are marked with an double asterisk (**) next to the Cite No. are not supplied because they were previously cited by or submitted to the Office in a prior application relied upon in this application for an earlier filing date under 35 U.S.C. 120.

[1]Applicant's unique citation designation number (optional). [2]Applicant is to place a check mark here if English language Translation is attached.

Examiner Signature	PATRICIA A. DUFFY	Date Considered	6/8/04

Notice of References Cited

Application/Control No.	Applicant(s)/Patent Under Reexamination	
09/880,748	RUBEN ET AL.	
Examiner	Art Unit	Page 1 of 1
Patricia A. Duffy	1645	

U.S. PATENT DOCUMENTS

*		Document Number Country Code-Number-Kind Code	Date MM-YYYY	Name	Classification
	A	US-			
	B	US-			
	C	US-			
	D	US-			
	E	US-			
	F	US-			
	G	US-			
	H	US-			
	I	US-			
	J	US-			
	K	US-			
	L	US-			
	M	US-			

FOREIGN PATENT DOCUMENTS

*		Document Number Country Code-Number-Kind Code	Date MM-YYYY	Country	Name	Classification
	N					
	O					
	P					
	Q					
	R					
	S					
	T					

NON-PATENT DOCUMENTS

*		Include as applicable: Author, Title Date, Publisher, Edition or Volume, Pertinent Pages)
	U	Kennell, D.E., Prog. Nucl. Acid Res. Mol. Biol., 11: 259-301, 1971
	V	
	W	
	X	

*A copy of this reference is not being furnished with this Office action. (See MPEP § 707.05(a).)
Dates in MM-YYYY format are publication dates. Classifications may be US or foreign.

Principles and Practices of Nucleic Acid Hybridization

DAVID E. KENNELL

Department of Microbiology, Washington University School of Medicine
St. Louis, Missouri

I. Introduction 259
II. Yield of Hybrid as a Function of Reactant Concentration . 262
A. The Law of Mass Action 264
B. The Reaction Is Reversible 267
C. Estimation of Equilibrium Constant for RNA-DNA Reaction 271
III. Kinetics of the Hybridization Reaction 275
A. Second-Order Kinetics and Estimation of Genome Heterogeneity 275
B. First-Order Kinetics and the Reaction of RNA with DNA 276
IV. Estimating Relative Transcribing Activities of DNA Sites . 281
A. A Simple Case: Fraction of Genome for rRNA . . 281
B. A Complex Case: Total Cell RNA 282
V. Quantitative Limitations of Hybridization 283
A. Kinetic Considerations 284
B. Thermodynamic Considerations 284
VI. Use of Competitor RNA to Estimate Specificity . . . 286
A. Some Basic Considerations 287
B. A Not So Simple Case: Estimating the Fraction of Pulse-Labeled RNA That Is rRNA 287
C. A Very Complex Case: Estimating Relatedness of RNA Populations from Different Cells 288
VII. Hybridization of Nucleic Acids from Eucaryotic Cells . . 292
A. Factors Adding to the Complexity of Interpretations . 292
B. Stringency of Reaction Conditions and Thermal Stabilities of Complex 293
C. Expected Reaction Kinetics for Hybridization of Mouse RNA to DNA 294
D. The Reaction Is Limited for Thermodynamic and Kinetic Reasons 296
E. Presaturation Competition 297
VIII. Conclusions 298
References 299

I. Introduction

The formation of stable complexes of RNA with denatured strands of complementary DNA (*1*, *2*) provides a biochemical assay that relates

DNA to the immediate product of gene expression. In the ten years since the initial observation, RNA·DNA hybrid formation has been an important part of many thousands of research papers. Probably no other assay has found such widespread use in molecular biology.

Since the chemistry of RNA·DNA hybrid formation is similar to that of DNA·DNA renaturation, most of its basic aspects are common to the DNA·DNA reaction. The chemistry of denaturation and reassociation of DNA are discussed in the extensive reviews by Marmur, Rownd, and Schildkraut (*3*) and Felsenfeld and Miles (*4*). Most chemists still believe that the major forces holding together a double-stranded nucleic acid are derived from the cooperative "stacking-free energy" (*3, 5*). Hydrogen bonds, while probably important for the specificity of the reaction, are thought to play a minor part in the total stability of the duplex in solution. In any case, many physical and chemical agents cause strand separation or denaturation of DNA (*3*). The two most commonly used are heat and high pH. However, many organic molecules are also effective. Formamide, in particular, has been useful (*6–8*) because at high concentrations it causes denaturation at room temperature without intrastrand scissions (*6, 7*). Although the rate of breakage of phosphodiester or glycosyl bonds in RNA is relatively slow at the usual annealing temperatures (*9*), such breakage has been observed with ribosomal RNA (*7, 10*). Thus, formamide is useful in studies in which recovery of the intact hybrid RNA is desired or when very long annealing times are required.

Many factors affect the reassociation of denatured DNA. In general, the optimal temperature of reassociation (T_r) is related to the melting temperature (T_m) and is about 25° lower (*3, 11*). The rate of renaturation is dependent on several factors (3) including salt concentration, G + C content, and the viscosity of the solvent (*12, 13*). Wetmur and Davidson (*14*) showed that the rate was proportional to the square root of the strand length; longer strands react more rapidly than shorter ones. Furthermore, the rate of renaturation obeys second-order reaction kinetics (*3, 14*); this means that the rate of reaction is a function of the reactant concentrations. It follows that strands from a complex genome will renature more slowly than those from a less complex genome, given that all other parameters, such as size, total concentration, G + C content, etc. are the same (*15*).

The direct relationship between the product of concentration of complementary strands to the rate of reassociation suggested that the initial rate-limiting step in reassociation is a nucleation event caused by the collision and formation of a few correct base pairs (*3, 14*). The subsequent "zippering" reaction is presumably fast and results in the ultimate formation of stable complex. However, there is some minimal length below which the duplex will not be stable. Various experiments have

ession. In the ten years since
formation has been an im-
h papers. Probably no other
olecular biology.
d formation is similar to that
c aspects are common to the
turation and reassociation of
vs by Marmur, Rownd, and
). Most chemists still believe
louble-stranded nucleic acid
ee energy" (*3, 5*). Hydrogen
ecificity of the reaction, are
lity of the duplex in solution.
ents cause strand separation
commonly used are heat and
re also effective. Formamide,
se at high concentrations it
vithout intrastrand scissions
phodiester or glycosyl bonds
ling temperatures (*9*), such
RNA (*7, 10*). Thus, form-
of the intact hybrid RNA is
ire required.
lenatured DNA. In general,
r) is related to the melting
, *11*). The rate of renatura-
ncluding salt concentration,
vent (*12, 13*). Wetmur and
portional to the square root
e rapidly than shorter ones.
cond-order reaction kinetics
is a function of the reactant
a complex genome will re-
omplex genome, given that
centration, G + C content,

ict of concentration of com-
n suggested that the initial
ation event caused by the
pairs (*3, 14*). The subse-
and results in the ultimate
e is some minimal length
Various experiments have

been performed with synthetic polynucleotides (*16–19, 19a*) or with natural nucleic acids that had been sheared to defined sizes (*20, 21*) in order to estimate this minimal length. These studies have been reviewed by Walker (*22*) and Thomas (*23*), and it would appear that, depending on G + C content (*11, 21a, 24*), the minimum size for a stable complex is from 10 to 20 nucleotides. The thermal stability rises sharply for longer lengths so that, depending on the G + C content (*21a*), the stability of a complementary duplex of 25–50 nucleotides approaches that of any much longer complex (*15, 18, 20, 21a*).

The thermal stability of a nucleic acid duplex is extremely sensitive to the presence of mismatched nucleotide pairs within the polymer strands (*25–27, 27a*). The reduced stabilities of many associated complexes from cells of higher organisms results from variable extents of mismatching within duplexes in the hybrid population (*15, 20, 22, 27*). There seems to be a need for more quantitative studies with defined synthetic polymers, to determine the effects on thermal stabilities of defined types of mismatching within the complex. For example, it is not clear whether mismatching normally occurs in clusters of contiguous bases or whether the noncomplementary bases are scattered at random throughout the molecule.

Many studies have been possible only because of technical improvements in the procedures of hybridization. These have been reviewed by Gillespie (*28*). The use of agar as a supporting medium for the denatured DNA (*29*) was widely used until introduction of nitrocellulose filters for trapping the DNA (*30*). The filter method has many advantages with respect to simplicity, reproducibility, and relatively low nonspecific background sticking. However, there are certain hybridization studies, such as those involving high RNA/DNA inputs, for which it may be advantageous to react in solution before trapping the hybrid on a filter (*31*).

Single-stranded DNA and RNA can be separated from duplex structures by chromatography on calcium phosphate (hydroxylapatite) columns (*21a, 32–34*). This supporting medium is useful for many kinds of studies since it is stable over a wide range of temperatures and to organic solvents. Furthermore, it provides an instrument for analysis or purification of preparative amounts of nucleic acids (*15, 35, 35a*). It is also useful for studying oligonucleotide-polynucleotide complexes (*19a*).

Progress in another area of biology involving hybridization has been of great significance for study of gene expression. It has become possible to study RNA synthesized from one or a few genes by use of DNA carrying only a small fraction of the genome. So far, this last development has been of great value for an analysis of the expression of the genes of phage λ (*36*) and the *E. coli* genes responsible for tryptophan synthesis (*37, 38*). Limited segments of the bacterial chromosome can be carried

either on episomal F factors (*39*) or as a part of a transducing bacteriophage. Transducing phages are much better suited than episomal carriers for purification of preparative amounts of DNA that can be enriched 100-fold or more for selected genes of the bacterial chromosome. As might be predicted there is a high degree of background homology between either an episomal factor (*39*) or a generalized transducing phage (*40*) and total *E. coli* RNA. Thus, specialized transducing phages have been the most useful. Signer and Beckwith transposed the *lac* operon close enough to the chromosomal attachment site of λ or *att* φ80 to allow their transduction (*41*) and procedures have been outlined to construct such strains with any selectable marker (*41*, *42*).

For any assay to be useful it must be quantitative. In the next sections some features of the hybridization reaction that bear on quantitative interpretations are discussed. Technical considerations such as optimal salt, temperature, volume, etc., have been reviewed (*28*, *30*, *31*). Rather than dwell on these, it might be useful to discuss the factors that determine the yield in any complex reaction that is the summation of several hundred reactions. Such is the case in the reaction of RNA from a cell to its DNA; however, many of the following comments could apply to any multireaction system.

II. Yield of Hybrid as a Function of Reactant Concentration

The complementarity of an RNA to a given DNA is monitored in the hybrid reaction by the percentage of the reacted RNA that forms a stable hybrid. Ideally, at a given RNA/DNA input, all RNA molecules for which excess complementary DNA sequences are available should be bound at equilibrium. Unfortunately, a fraction of the RNA for which complementary DNA sites are available may not be bound. This is shown most easily by observing the fraction of RNA bound when the concentration of RNA and DNA are varied but their *ratio* is kept constant. In this situation, if complementarity were the only variable, the fraction bound should remain constant. This is clearly not the case (*31*) (Fig. 1).

The concentration curves show that the equilibrium constant, K, for the reaction: RNA + DNA $\rightleftarrows$ RNA·DNA is such that at sufficiently low concentrations of reactants, a significant fraction of the pulse-labeled RNA is not bound at equilibrium, even though sites are available. However, if the label is only in the stable RNA molecules that themselves represent about 97% of the total RNA of *E. coli*, then the yield is not concentration-dependent at similar RNA/DNA inputs (*31*, *43*).

These differences between RNA labeled only in stable components and RNA that is pulse-labeled reflect the very different distributions of isotope

art of a transducing bacterio-
suited than episomal carriers
DNA that can be enriched
ie bacterial chromosome. As
of background homology be-
eneralized transducing phage
zed transducing phages have
h transposed the *lac* operon
it site of λ or *att* φ80 to allow
ve been outlined to construct
42).

ntitative. In the next sections
iat bear on quantitative inter-
rations such as optimal salt,
ed (*28, 30, 31*). Rather than
ss the factors that determine
e summation of several hun-
ion of RNA from a cell to its
mments could apply to any

eactant Concentration

en DNA is monitored in the
ted RNA that forms a stable
put, all RNA molecules for
ces are available should be
tion of the RNA for which
not be bound. This is shown
bound when the concentra-
atio is kept constant. In this
variable, the fraction bound
the case (*31*) (Fig. 1).
equilibrium constant, *K*, for
such that at sufficiently low
on of the pulse-labeled RNA
s are available. However, if
s that themselves represent
ie yield is not concentration-
43).

ly in stable components and
erent distributions of isotope

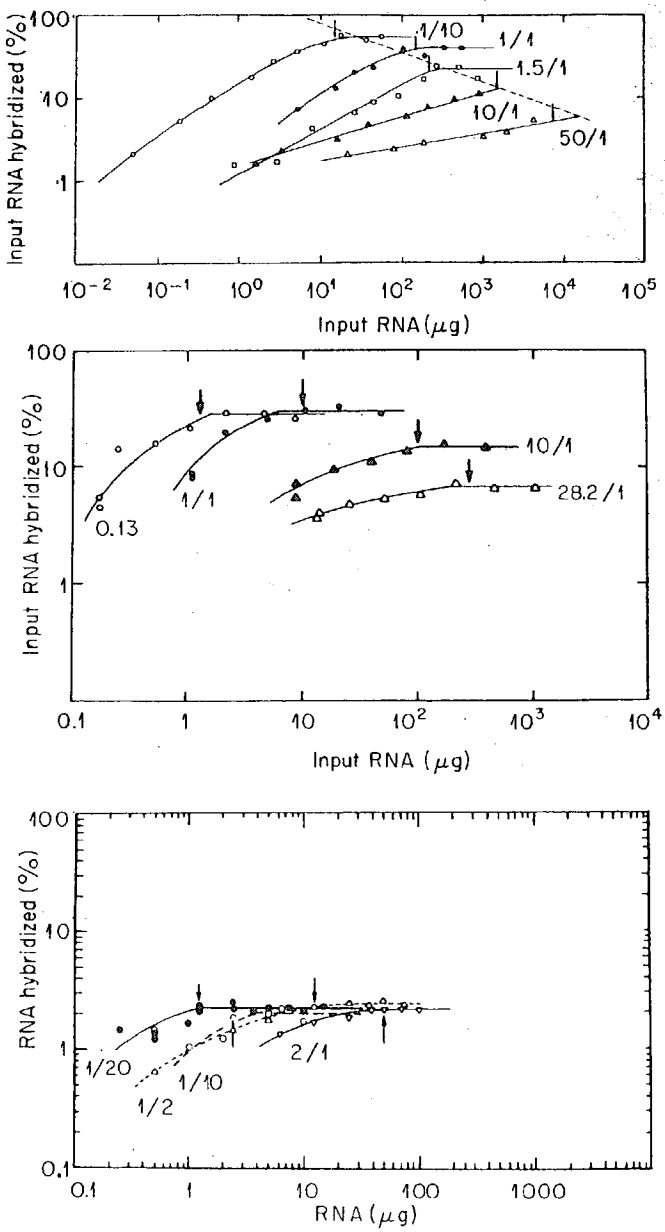


FIG. 1. Hybridization efficiency as a function of nucleic acid concentration. Pulse-labeled RNA in 0.5 ml of 0.9 M NaCl + 0.09 M Na citrate was allowed to react for 20 hours at 66°C with denatured DNA. The μg RNA/μg DNA inputs, indicated

in the same RNA molecules of the two preparations. In pulse-labeled *E. coli* RNA, about 50% of the isotope is in the mRNA fraction, but the mRNA comprises only 3% of the total RNA (*43*). Furthermore, the isotope in this minor fraction of the RNA mass is distributed among several hundred different species, each of which is complementary to a unique DNA locus. The numbers of molecules in each species vary over at least a 1000-fold range (*43, 44*). Thus, at any given RNA/DNA input, between the limits of all DNA sites filled to all being in excess, certain DNA sites are filled, others may be virtually empty, while still others are partially filled. At a given RNA/DNA input, this latter class contributes to the "active" titration of the DNA and is the major contributor to a concentration dependence of the reaction. However, all species, including those for which there are more copies than there are complementary DNA sites, may not be fully bound; this is true at sufficiently low concentrations of reactants (see below).

Another way to show a concentration dependence is to measure the hybrid yield when the amounts of RNA and DNA as well as salt concentrations are kept constant but the volume is changed. There can be a marked increase in fraction hybridized when the volume is decreased (concentration is increased) (*31, 45*).

A. The Law of Mass Action

These relationships are shown most clearly by examination of the bimolecular association reaction between one RNA species R, and its DNA site, D, with the moles bound at equilibrium, y. Let R and D represent the initial concentration of reactants. The concentrations of bound versus free reactants at equilibrium is given by:

$$\frac{(R - y)(D - y)}{y} = K \tag{1}$$

or

$$(R)(D) - (D + R + K)y + y^2 = 0 \tag{2}$$

For a given value of K, the dissociation constant, values of R and D determine y from the quadratic equation above. The value of y/R

by the ratios, are the same for all points on a given curve. *Top: Escherichia coli* RNA with *E. coli* DNA trapped on nitrocellulose filters. The arrows indicate 150 μg *E. coli* DNA. *Middle:* RNA from T4-infected *E. coli*, labeled between 21 and 25 minutes after infection reacting with T4 DNA free in solution. The arrows indicate 10 μg DNA. *Bottom:* RNA from *E. coli* fully induced for *lac* operon RNA. The arrows indicate 25 μg of $\lambda d1ac$ DNA trapped on nitrocellulose filters. The top and middle frames are from Kennell and Kotoulas (*31*).

preparations. In pulse-labeled
the mRNA fraction, but the
43). Furthermore, the isotope
istributed among several hun-
plementary to a unique DNA
ecies vary over at least a 1000-
NA/DNA input, between the
excess, certain DNA sites are
still others are partially filled.
ss contributes to the "active"
ributor to a concentration de-
ies, including those for which
ementary DNA sites, may not
w concentrations of reactants

ependence is to measure the
d DNA as well as salt con-
e is changed. There can be a
en the volume is decreased

early by examination of the
ne RNA species R, and its
librium, y. Let R and D rep-
The concentrations of bound
by:

$$\frac{}{-y)} = K \qquad (1)$$

$$+ y^2 = 0 \qquad (2)$$

onstant, values of R and D
above. The value of y/R

curve. *Top: Escherichia coli* RNA
he arrows indicate 150 μg *E. coli*
led between 21 and 25 minutes
tion. The arrows indicate 10 μg
lac operon RNA. The arrows in-
lose filters. The top and middle

(fraction of RNA hybridized) versus R for a given R/D will give a concentration-dependence curve. Such families of curves for different R/D inputs are most easily generated by using a computer to solve the quadratic equation for different values of R, D and K. One such family is shown in Fig. 2 for the arbitrarily chosen value of $K = 10^{-4}$.

Each of the four curves represents binding of one of four species of RNA (R_1 through R_4) and each species is ten times more concentrated than a preceding species; this gives a 1000-fold range of concentrations. In all cases, the slopes approach a value of 1 at very low inputs of DNA and RNA; i.e., the fractional yield becomes proportional to the amount of DNA reacted. This follows from rearrangement of Eq. (1) to give

$$\frac{y}{R} = \frac{D}{K} - \frac{yD}{KR} - \frac{y}{K} + \frac{y^2}{KR} \qquad (3)$$

For $R \gg y \ll D$, $y/R \to D/K$. The fraction bound is proportional to the amount of DNA reacted when the fraction of RNA hybridized is a small proportion of the maximum that could be hybridized at that input ratio.

As noted above, even that species in excess with respect to its DNA site (RNA/DNA input = 10/1 curve) is not maximally bound at equilibrium under conditions of sufficiently low input concentrations. However, this species is always more fully bound (y/R is closer to its maximum value) than are the other species, except at very high inputs of RNA and DNA, when the binding of all species is maximal. For example, at $D = 10^{-3}$, virtually all of R_1 for which sites are available (10%) is bound, while R_3 and R_4 are about 90% and R_2 about 73% bound. At $D = 10^{-4}$, the

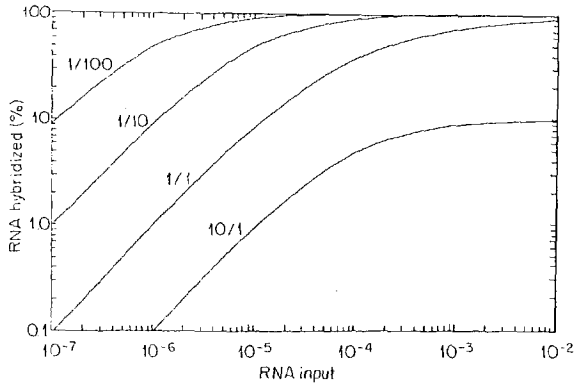


FIG. 2. Expected curves for the percentage of input RNA hybridized as a function of RNA concentration at the indicated RNA/DNA ratio. Each curve is for one species of RNA to its DNA site when the equilibrium constant, K, is 10^{-4}.

binding of R_1 is 90% of maximum while R_3 and R_4 are 49% and R_2 is only 38%. At $D = 10^{-5}$, the same values are 49%, 9%, and 8.5%, respectively. In terms of the total RNA hybrid ($\Sigma y/R$), it can be seen that in this range of R/D inputs its variation with total RNA concentration ($R_1 + R_2 + R_3 + R_4$) is mainly a function of the R_2 species (R/D input = 1). When all species are fully bound, R_3 and R_4 contribute only 10% and 1%, respectively, as much hybrid as does either R_1 or R_2. Thus, in general, if the hybrid binding is submaximal, it is so mainly due to the incomplete binding of the RNA molecules that are just filling their sites at the given RNA/DNA input, i.e., for which the RNA/DNA input is in the range of 1.

One conclusion from this kind of analysis is that as the RNA/DNA input is increased and more and more DNA sites are filled, the slopes of the concentration-dependence curves should become less steep. The most simple way to visualize this is to consider the sequential filling of the DNA at the RNA/DNA inputs at which (a), the most abundant mRNA's are just filling their sites and thus contributing a large fraction of the hybridized label, compared to (b) the least abundant mRNA's are just filling their sites and thus contributing only a small fraction of the labeled hybrid. In the latter case, most of the DNA sites that can be filled would be saturated at lower RNA/DNA inputs. Therefore, the percentage contribution to the total hybrid by the more rare RNA species is not as great when they fill their sites as was that of the more abundant molecules at the RNA/DNA inputs when they just filled their sites. Since the major contributors to a concentration dependence are those molecules that are just filling their sites, the dependence at high RNA/DNA input, when the rare species are being bound, is not as obvious experimentally. Thus, the slopes of the curves are less steep; this relationship is seen in Fig. 1.

The major contributors to a concentration dependence on yield are those species that are just filling their DNA sites, i.e., they are in the same range of concentrations as are their DNA sites. Thus, in Fig. 2 at $D = 10^{-3}$ the fraction bound (y/R) is about 75% of the maximum possible at that R/D input with at least 95% of the unbound RNA, for which DNA sites are available, being species R_2 ($R_2/D_2 = 1$). Suppose the total RNA input ($R_1 + R_2 + R_3 + R_4$) per unit DNA were increased 10-fold. This would now mean that each RNA species represented by a curve in Fig. 2 would be represented by the adjacent curve on the right, e.g., that RNA species that had been 1/10 input would now be 1/1 input and this species would now provide the major contribution to a concentration dependence on yield. Again the amount of total DNA required to give a maximal amount of hybrid would be about 10^{-2}, i.e., the fractional yield would be proportioned to the same amount of DNA. Thus, in a heterogeneous

mixture of RNA molecules in which there is a fairly continuous distribution of numbers per species, the amount of DNA required to give a maximum yield of hybrid is the same at all RNA/DNA inputs. In reexamining the concentration-dependence curves for pulse-labeled *E. coli* and T4 RNA, it can be seen that this rule is correct within the limitations of these measurements; about 150 μg of DNA is required for maximum efficiency in the *E. coli* system and only 10 μg in the T4 hybridizations (Fig. 1). There are obvious exceptions to this rule, e.g., if one is measuring radioactivity and the specific activities of very rare species far exceed those of the other molecules.

Also, it should be emphasized that the amount of DNA required for maximal efficiency will depend on the heterogeneity of the DNA and the reaction conditions. For the reaction of *E. coli* RNA to its DNA and T4 RNA to its DNA, in the usual ranges of DNA reacted (1 μg to 100 μg), we observed that the hybrid yields were less dependent on concentrations at a given RNA/DNA input when both RNA and DNA were completely free in solution than when DNA was trapped in a filter (*31*). This suggests that the association constant for the reaction is higher when both species are in solution. It is probable that when the reaction occurs in a filter, the DNA is constrained to migrate within the cavities of the filter. I have assumed in this and the following section (III) that the reaction can still be treated by the same thermodynamic and kinetic considerations as for the reaction in solution, although presumably the activity coefficients of the nucleic acids must be different in the two cases.

Concentration dependence on yield would be less serious if reaction conditions were modified so as to increase the association constant. However, it is not obvious that this could be done to a significant extent without also decreasing the specificity of the reaction. For example, complex formation is favored at lower temperatures. However, while the yield of complex can be increased by reaction at low temperatures (*46*), e.g., 40°, the specificity as well as the thermal stabilities of the resulting duplexes is greatly reduced (*20, 45–47*).

It should be emphasized that the concentration dependency for final fractional yield of hybrid has no direct relationship to the kinetics of the reaction. As is discussed in Section III, the reaction of bacterial or bacteriophage nucleic acid can be complete at 16 hours. The curves of Fig. 1 show yields at equilibrium.

B. The Reaction Is Reversible

Another way of describing a concentration dependency for yield is simply to state that the reaction can be reversed. If true, there should

be another way to show this reversibility; hybrid RNA should be *lost* from DNA to an extent that is also concentration dependent. This can be shown using Eq. (1). Assume that an RNA reacts with DNA on a filter to give hybrid y. After washing, treating with RNase, and again washing, the filter with hybrid is placed into a fresh volume of solution and incubated at the usual hybridization temperature. In general, some of y will be composed of DNA sites that are filled, while the remainder will be made up of DNA species filled to variable extents. First, consider the filled sites. For these, $R - y = D - y$ at any time of the second incubation. Therefore, $(R - y)^2/y = K$, so that increasing y 100-fold would increase $R - y$ by only 10-fold and the fraction of free RNA $(R - y)$ per fraction bound (y) would decrease by 10-fold; in general, this ratio would be inversely proportional to $y^{1/2}$.

The second class of hybrids would include those DNA sites only partially filled. For many of these, $D \gg y$ so that $(R - y)/y = K/D$. An increase of y by 100-fold could only come from a large increase of D (perhaps not 100-fold), since $D - y \simeq D$. If D were also 100-fold higher, this would decrease to 0.01 the equilibrium value of free RNA $(R - y)$ per bound RNA (y). For intermediate values between DNA sites completely filled to DNA sites slightly filled, the ratio of bound to free RNA will vary between $y^{1/2}$ and y. Since the ratio of sites filled to sites partially filled is a function of the RNA/DNA input, the exact relationship will be different at each different input ratio.

This predicted result is shown experimentally in Fig. 3. *E. coli* [^{3}H] RNA was reacted to form hybrid with 1 μg or 100 μg *E. coli* DNA filters. The RNA/DNA inputs were chosen arbitrarily at 1/2 in both cases. Since the fraction of RNA hybridized is lower with 1 μg DNA, the difference between y in the two cases is increased considerably more than 100-fold. It can be seen that a much larger fraction of RNA hybrid is lost upon reincubation of the 1 μg DNA filters than is lost from the 100 μg DNA filters.

Since a larger fraction of RNA is lost from the 1 μg DNA than from 100 μg DNA hybrid at a given temperature, the melting curve of the former should be shifted to lower temperatures. This is shown in Fig. 3. The T_m for elution in 0.15 M NaCl + 0.015 M Na citrate for hybrid from 100 μg DNA filters is 87°, and for 1 μg DNA filters it is 79°. Since both curves are very steep, and equally so, homogeneity with respect to intrinsic stability of the complexes is indicated. However, this result should be viewed with reservation since preliminary results suggest that there is damage to the filters at 90° and part or all of this difference could be explained by a larger fractional loss of the DNA, itself, from 1 μg than from 100 μg DNA filters.

y; hybrid RNA should be *lost*
:entration dependent. This can
n RNA reacts with DNA on a
eating with RNase, and again
into a fresh volume of solution
temperature. In general, some
are filled, while the remainder
variable extents. First, consider
y at any time of the second
so that increasing y 100-fold
the fraction of free RNA ($R -$
ıse by 10-fold; in general, this
$y^{1/2}$.
include those DNA sites only
so that $(R - y)/y = K/D$. An
ne from a large increase of D
If D were also 100-fold higher,
m value of free RNA ($R - y$)
alues between DNA sites com-
he ratio of bound to free RNA
io of sites filled to sites partially
ɔut, the exact relationship will

ıentally in Fig. 3. *E. coli* [^{3}H]
g or 100 μg *E. coli* DNA filters.
arily at 1/2 in both cases. Since
with 1 μg DNA, the difference
ɔnsiderably more than 100-fold.
ɔn of RNA hybrid is lost upon
ı is lost from the 100 μg DNA

from the 1 μg DNA than from
ture, the melting curve of the
atures. This is shown in Fig. 3.
5 M Na citrate for hybrid from
NA filters it is 79°. Since both
homogeneity with respect to
dicated. However, this result
reliminary results suggest that
part or all of this difference
loss of the DNA, itself, from

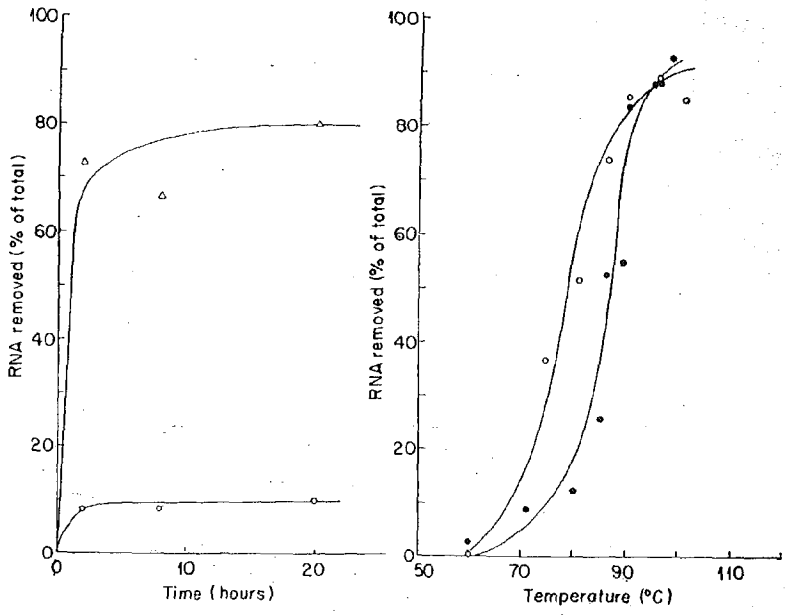


FIG. 3. Loss of pulse-labeled *Escherichia coli* RNA from preformed *E. coli* RNA·DNA hybrid (RNA/DNA input of ½) as a function of DNA concentration. *Left:* Incubation of filter in 0.5 ml 0.9 M NaCl + 0.09 M Na citrate at 66°C (○) 100 μg DNA, (△) 1 μg DNA. *Right:* Incubation in 0.5 ml 0.15 M NaCl + 0.015 M Na citrate for 15 minutes at the indicated temperature. After the final incubations the filters were removed and washed, treated with RNase, and again washed before counting (*31*). (●), 100 μg DNA; (○), 1 μg DNA.

Gillespie and Spiegelman (*30*) observed complete retention of DNA to the filters after 20 hours of hybridization at 66°. We have not observed losses either, but a few reports have claimed small but measurable losses of DNA in 24 hours at 66° (3–8%) (*48*). One group claimed 28 and 41% losses of pea DNA (*7*). It would be difficult to prove by measuring loss of labeled DNA that the differences between 1 μg and 100 μg DNA filters are not due to loss of the entire DNA·RNA since a very small fraction (<5%) of the DNA is complexed; perhaps only the unbound DNA is retained.

However, another experiment was designed to show that the DNA remains competent during the 20 hours of reincubation. If the reaction were reversible, it should be possible to show exchange when [^{3}H]RNA·DNA hybrids are reincubated with [^{14}C]RNA. Furthermore, according to Eq. (1), the extent of exchange (percentage) should be greater for lesser amounts of hybrid.

E. coli filters containing 1 μg and 100 μg of DNA were allowed to react with *E. coli* [^{3}H]RNA, with label only in stable components, at RNA/DNA inputs of 1/20. Ribosomal RNA is in about 10-fold excess over its DNA sites at this ratio. After 20 hours, the filters were removed, washed, and then reincubated with *E. coli* [^{14}C-stable]RNA at RNA/DNA inputs of 1/20. The losses of [^{3}H]RNA from the 1 μg and 100 μg DNA filters were 60% and 18%, respectively, in the second 20-hour incubation.

In the second incubation 24% of the initial [^{3}H]RNA hybrid on the 1 μg DNA filters was replaced by [^{14}C]RNA hybrid (44% recovery of lost [^{3}H]RNA); on the 100 μg DNA filters [^{14}C]RNA replaced 12% of the initial ^{3}H-hybrid to give 64% recovery of lost ^{3}H (Fig. 4). The incomplete recovery of lost [^{3}H]RNA may result from losses of some DNA sites from the filter or from some upper limit of efficiency of hybridization; for an initial annealing, it is 60 to 70% (*43, 49*). The important result is that the fractional yield of ^{14}C-hybrid is greater on the 1 μg than it is on the 100 μg DNA filters. This shows that at least part of the *increased* loss of

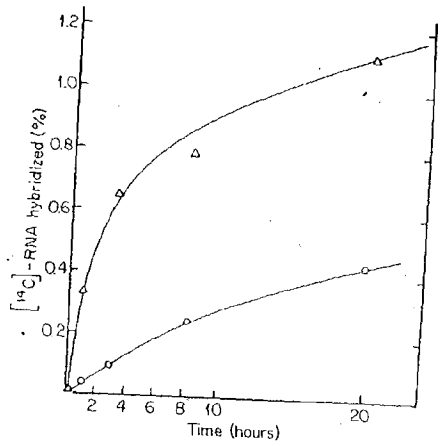
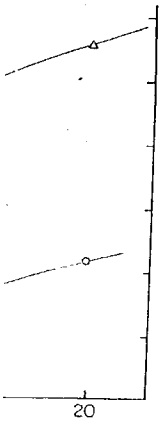


FIG. 4. The formation of *Escherichia coli* [^{14}C]RNA hybrids of stable RNA and DNA. These sites are saturated with [^{3}H]RNA at time zero; thus, the reaction shows exchange as a function of DNA concentration. [^{3}H]RNA, labeled only in stable RNA (*31*), was allowed to react with 1 μg (△) or 100 μg (○) DNA on a filter for 20 hours at an input RNA/DNA of 1/20. The filters were removed and washed on both sides (no RNase) and each reincubated in a fresh vial containing [^{14}C]RNA, labeled only in stable RNA, at an RNA/DNA input of 1/20. Each filter was removed at an indicated time, washed, treated with RNase, and again washed before counting. The zero-time filters (end of first incubation) had approximately 5% of the [^{3}H]RNA bound after the wash-RNase-wash treatment [1 μg had as high a fraction as 100 μg DNA, since there is no concentration dependence for stable RNA at this input ratio (*31*)].

g of DNA were allowed to react
in stable components, at RNA/
in about 10-fold excess over its
e filters were removed, washed,
able]RNA at RNA/DNA inputs
e 1 μg and 100 μg DNA filters
ond 20-hour incubation.
initial [^{3}H]RNA hybrid on the
NA hybrid (44% recovery of lost
[^{14}C]RNA replaced 12% of the
st ^{3}H (Fig. 4). The incomplete
losses of some DNA sites from
ciency of hybridization; for an
The important result is that the
the 1 μg than it is on the 100
part of the *increased* loss of



C]RNA hybrids of stable RNA and
time zero; thus, the reaction shows
]RNA, labeled only in stable RNA
0 μg (○) DNA on a filter for 20
were removed and washed on both
vial containing [^{14}C]RNA, labeled
20. Each filter was removed at an
gain washed before counting. The
ximately 5% of the [^{3}H]RNA bound
s high a fraction as 100 μg DNA,
ble RNA at this input ratio (*31*)].

[^{3}H]RNA from 1 μg filters is from DNA sites still competent to form new hybrid.

These results support the conclusion that the reaction of RNA with DNA obeys the law of mass action, and as a result, the fractional yield of hybrid as well as the loss of preformed hybrid are concentration-dependent.

C. Estimation of Equilibrium Constant for RNA-DNA Reaction

For the simple two-component system considered in Fig. 2, the fraction of RNA bound [y/R at any given concentration of RNA (R) and DNA (D)] determines a unique value of K. However, because of the relative crudeness of the measurements, K is measured most accurately at reactant concentrations sufficient to hybridize a significant fraction of RNA but low enough so that the maximal fraction has not been hybridized, e.g., so that 50% of the RNA is bound rather than 1% or 99%.

The *E. coli* and T4 curves in Fig. 1 show a concentration dependence for the hybridization of whole cell RNA to its DNA; as such, they give the total hybrid from the reaction of several hundred species of RNA, each containing a unique number of molecules. As indicated in the preceding discussion, almost all of the concentration dependency is due to the RNA molecules that are just filling their DNA sites. It would be difficult to estimate a value of K for the reaction since the more abundant RNA species would provide an unknown background of bound label. However, it should be possible to make a rough estimate of K for the reaction of RNA from the lactose operon (*lac* RNA) to *lac* DNA (Fig. 1). Of necessity, it would be very approximate for a number of obvious reasons, including the fact that any bacterial mRNA species is composed of molecules that vary from a very small size (perhaps a trinucleotide) to a maximum corresponding to the length of the operon (*50*). However, it may be of some value to make these calculations without taking the exact answers too seriously.

Lac RNA is synthesized as a polycistronic mRNA from the three genes Z, Y, and A (*51*). The length of the Z gene is 3700 nucleotides, (*52, 53*), the Y gene is probably about 900 nucleotides in length, judging from recent estimates of the permease size (*54*), while the A gene is also about 900, judging from the size of β-galactosidase transacetylase (*55*). The O gene is very small ($<$100 nucleotide pairs) (*56*). This gives a total length of about 5500 nucleotides in a double strand of DNA of M.W. $(330)(2)(5500) = 3.6 \times 10^6$.

Lac mRNA was detected by annealing it to *lac* DNA carried in the

defective λ phage, λ*dlac*. The fraction of the λ*dlac* DNA that is *lac* operon can be estimated from the difference in densities of λ and λ*dlac*; the normal phage has a $\rho = 1.495$ while its defective counterpart has a $\rho = 1.502$. Normal λ has a DNA/protein ratio of one, and its DNA has a $\rho = 1.7093$ (*57*, *58*). Since the resultant density of the phage is close to the sum of the density contribution from the DNA and protein (*59*), the protein has a $\rho = 1.28$. If we assume, as did Weigle *et al.* (*59*), that the increments of mass and volume in the defective phage are due entirely to the increment of DNA and that this DNA has the same density as the original DNA, then the change in phage density is

$$\Delta\rho = \rho_2 - \rho_1 = \frac{P + D + \alpha D}{V + \alpha D/\rho_D} - \rho_1 \qquad (4)$$

where P and D are, respectively, the masses of protein and DNA per volume V of the normal phage. α is the fractional increment of DNA, and ρ_1 and ρ_D are the densities of the normal phage and the DNA, respectively. When $V = 1$ ml,

$$P + D = \rho_1 \quad \text{and} \quad \alpha D = \frac{\rho_D \Delta_\rho}{\rho_1 - \rho_D + \Delta_\rho} \qquad (5)$$

Substitution of appropriate values into Eq. (5) gives $\alpha D = 0.060$; i.e., the defective phage has 6% more DNA than λ. Since the M.W. of λDNA is 30×10^6 (*60*), that of the λ*dlac* DNA is 31.8×10^6. Thus the ZYA segment is 11.3% of this λ*dlac* DNA.

RNA was purified from *lac*-induced bacteria and allowed to react with λ*dlac* DNA. Hybrid yields were measured over a wide range of nucleic acid concentrations at each of a number of RNA/DNA inputs (Fig. 1). In this system, maximum yields were obtained with about 25 μg DNA. The equilibrium constant, K, can be estimated from any of the curves by Eq. (1). When y = half-maximum, $(R - y) = y$, and $K = D - y$. For example, when RNA/DNA = 2, the reaction is half-maximum with 2 μg DNA or $2 \times 0.113 \times 0.5 = 0.113$ μg of single-stranded *lac* DNA. Since mRNA is about 3% of the RNA mass and contains about 50% of the pulse-label (*43*), the *lac* RNA input = 0.0072 μg in a reaction volume of 0.5 ml with 0.0036 μg bound = y. Since the M.W. of *lac* DNA = 1.8×10^6, $D - y = K = (0.194\ \mu g/0.5\ ml)\ (10^3\ ml/l)\ (10^6\ gm/\mu g)\ [1/(1.8 \times 10^6)] = 1.2 \times 10^{-10}$ M.

Lavallé and DeHauwer (*61*) estimated K for reaction of *trp* mRNA to DNA from the defective φ80 phage, φ80 *dtrp*, which contains DNA from the tryptophan operon. In their procedure, an RNA sample was hybridized with two DNA-containing filters in succession to give $K = D/(y_1/y_2 - 1)$ (*61*). The results led to an estimated dissociation constant

e λ*dlac* DNA that is *lac* operon
nsities of λ and λ*dlac;* the nor-
ve counterpart has a $\rho = 1.502$.
and its DNA has a $\rho = 1.7093$
e phage is close to the sum of
d protein (*59*), the protein has
t al. (*59*), that the increments
e are due entirely to the incre-
e same density as the original

$$\frac{+ \alpha D}{D/\rho_D} - \rho_1 \qquad (4)$$

sses of protein and DNA per
ctional increment of DNA, and
phage and the DNA, respec-

$$\frac{\rho_D \Delta_\rho}{- \rho_D + \Delta_\rho} \qquad (5)$$

q. (5) gives $\alpha D = 0.060$; i.e.,
λ. Since the M.W. of λDNA
is 31.8×10^6. Thus the ZYA

eria and allowed to react with
over a wide range of nucleic
RNA/DNA inputs (Fig. 1).
ined with about 25 μg DNA.
ed from any of the curves by
$) = y$, and $K = D - y$. For
n is half-maximum with 2 μg
gle-stranded *lac* DNA. Since
ntains about 50% of the pulse-
n a reaction volume of 0.5 ml
W. of *lac* DNA $= 1.8 \times 10^6$,
1) (10^6 gm/μg) [1/(1.8 ×

K for reaction of *trp* mRNA
) *dtrp*, which contains DNA
edure, an RNA sample was
s in succession to give $K =$
timated dissociation constant

of about 10^{-10} M, which is very close to that derived here from concentration dependence curves.

Bishop recently estimated a dissociation constant from the reaction of rRNA to DNA (*48*). The estimated values were $K = 0.114$ μg ml^{-1} for a 24-hour reaction or 0.047 μg ml^{-1} after extrapolating to infinite time. These would give values for K of 0.76×10^{-10} M and 0.32×10^{-10} M, respectively, which are somewhat lower than the other two estimations. In contrast, for SPO1 phage RNA-DNA hybrids, Gage (*62*) estimated a K value of 5×10^{-10} M. This estimate was based on the assumption that the 50% value for RNA hybridized at certain concentrations of DNA excess reflected thermodynamic aspects of the reaction only. Since this is probably close to the maximum value achievable, it is undoubtedly a very rough approximation. However, tentatively, these crude estimations point to a dissociation constant for completely complementary RNA·DNA hybrid in the neighborhood of 10^{-10} M.

The dissociation constant, K, can be used to calculate a free energy change, ΔG, for the reaction at 66°C,

$$\Delta G = -2.3\, RT \log K = -2.3 \times 2.0 \times 339°K \log(1.2 \times 10^{-10}) = +15{,}600 \text{ cal.}$$

Thus, there are about 16 kcal/mole energy driving the reaction toward association. It is possible to make a rough estimate of the equivalent number of base-pairs formed to account for this energy. Two kinds of base pairings probably contribute most of the total free energy of duplex formation (*5, 65*) (see also references given in *65*). The first includes the initial base pairings, which is given by

$$\Delta G^\circ = -RT \ln \sigma$$

where σ is the equilibrium constant for the "nucleation" event. It has been estimated to be approximately 10^{-4} (*65*). The second bond is that involved in pairing of bases adjacent to one or more paired bases; its energy is influenced strongly by the stacking energy of the contiguous bases and can be given by

$$\Delta G^\circ = -RT \ln s^n$$

where n is the number of such paired bases. Thus, the total free energy is

$$\Delta G^\circ_{\text{TOTAL}} = -RT \ln (\sigma s^n)$$

and $K = \sigma s^n$.

The value of s at the reaction temperature (66°) can be estimated from the relation:

$$\frac{d \ln s}{dt} = \frac{\Delta H}{RT^2}$$

The enthalpy, ΔH, estimated from calorimetric data from the helix-coil transition of DNA, is close to -7 kcal/mole of base-pair (*63, 64*). At the melting temperature, T_m, $s = 1$. The melting curve for *lac* RNA in the salt concentration used in the reaction (0.9 M NaCl + .09 M Na citrate) is shown in Fig. 5 and suggests a T_m of 90°. Using these values and integrating between 339°K and 363°K gives a value of $s = 1.95$ at 66°. Using this value for s and the value of K derived above leads to an estimate of $n = 49$ base pairs.

As pointed out above, the *lac* RNA is made up of all sizes of molecules up to 5500 nucleotides. Probably the average size is no more than half this number, since about 75% of the molecules begin to decay in a bacterium before they are completed (*50*). Even so, 50 bases would represent only 2.0% of such half-lengths. This supports the notion that the polynucleotide duplex is in a dynamic state with regions of opening and closing occurring along its length and with the average net energy holding the complex together equal to the equivalent of energy in less than 100 nucleotide pairs. Such a model of opening-closing or "breathing" of DNA was suggested from rates of tritium exchange between the base-bound protons of the nucleotides and the solvent (*66*).

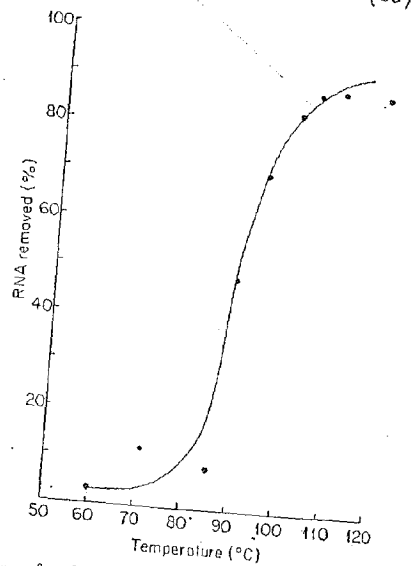


FIG. 5. Melting curve for *lac* mRNA-DNA hybrid on a filter. Each filter was placed in 0.5 ml of 0.9 M NaCl + 0.09 M Na citrate for 10 minutes at the temperature indicated. The filters were washed before counting.

netric data from the helix-coil
le of base-pair (*63, 64*). At the
ing curve for *lac* RNA in the
9 M NaCl + .09 M Na citrate)
0°. Using these values and in-
s a value of $s = 1.95$ at 66°.
derived above leads to an esti-

ade up of all sizes of molecules
age size is no more than half
ules begin to decay in a bac-
Even so, 50 bases would rep-
s supports the notion that the
e with regions of opening and
the average net energy hold-
ivalent of energy in less than
ning-closing or "breathing" of
exchange between the base-
solvent (*66*).

100 110 120
°C)

ybrid on a filter. Each filter was
ate for 10 minutes at the tempera-
unting.

III. Kinetics of the Hybridization Reaction

A. Second-Order Kinetics and Estimation of Genome Heterogeneity

In a simple bimolecular reaction

$$A + B \underset{k_2}{\overset{k_1}{\rightleftarrows}} AB$$

the rate of reaction is

$$dx/dt = k_1(a - x)(b - x) - k_2x \tag{6}$$

where a and b are the initial concentrations of reactants and x is the amount of complex at time (t). When $a = b$,

$$dx/dt = k_1(a - x)^2 - k_2x \tag{7}$$

If $k_1 \gg k_2$, the reaction proceeds far to the right and it is described as "irreversible." If this approximation is acceptable, $dx/dt = k_1(a - x)^2$, and integration and rearrangement gives the fraction bound as a function of time

$$\frac{x}{a} = \frac{ak_1t}{1 + ak_1t} \tag{8}$$

The time to complex one-half of a is

$$t_{1/2} = \frac{1}{k_1a} \tag{9}$$

The time to complex half the reactants is inversely proportional to the initial concentration. This relationship can be used to estimate genome heterogeneity when denatured DNA strands of approximate gene size are reassociated.

Let us assume that the DNA from one organism contains 1000 genes, each of which has a unique sequence, while the DNA of a second organism contains 10,000 unique sequences. Sequences in the first DNA will be ten times more concentrated than are the sequences in the more heterogeneous DNA. Since each segment reanneals independently of the others, the fraction of total DNA annealed in a given time should reflect this heterogeneity. Britten and Kohne (*15*) plotted the fraction of reassociated DNA (x/a) vs $a \cdot t$ for denatured DNA fragments from a variety of sources. The expected relationship was found when DNA from bacteria or phage was reannealed; the $t_{1/2}$ was proportional to the number of nucleotide pairs over several orders of magnitude. In contrast, DNA from eucaryotic cells contained a fraction that reannealed at dispropor-

tionately low values of $a \cdot t$. This suggested that such DNA contains many repeating sequences. For example, 40% of calf DNA was estimated to contain sequences that may be repeated 100,000 times while the remaining 60% of the DNA may contain only unique sequences (*15*).

These analyses rely on two assumptions regarding the nature of the reaction. First, the reaction is assumed to be second order (*3, 12–14, 67*) so that $t_{1/2}$ is inversely proportional to a in Eq. (9). Second, the reaction is assumed to be irreversible, i.e., k_2 is assumed to be zero in Eq. (6). While in these particular studies this assumption can be made without invalidating the broad conclusions, it should be recognized that it is not true that the only physical consideration is the collision rate (*15*), so that even very rare molecules would be annealed if one could only wait long enough. Of course, this consideration is crucial for the interpretation of many other experiments (Section VII).

If the reaction were not "irreversible," how would the unique sequences reanneal? In these studies the unique sequences were annealed after several days because several milligrams of DNA were used (*15*). Calf DNA has about 700 times more nucleotide pairs per cell than does *E. coli* DNA. At the concentration of 8600 μg calf DNA/ml used in these studies, the unique sequences were as concentrated as are the sequences in *E. coli* DNA at 17 μg/ml. This concentration should be adequate to give 100% binding at equilibrium. At the other extreme, the repetitive sequences form duplex structures with a reduced thermal stability (*15, 27, 46, 47*). However, even at only 2 μg calf DNA per milliliter, these sequences are at an enormous relative concentration (equivalent to 400 μg of *E. coli* DNA/ml if there are 100,000 copies of each sequence). Thus, the reaction is driven to the formation of complex. However, a significant fraction of the isolated complex dissociates very quickly (simply by washing) at the same temperature at which it was formed at the very high concentrations (*15, 27*). This demonstrates the obvious reversibility of the annealing reaction for these species. Of course, the reduced stability of these complexes implies that only a fraction (perhaps half) of the base pairs are complementary on the repetitious DNA segments that have the size used in these studies (*15*).

B. First-Order Kinetics and the Reaction of RNA with DNA

The kinetics of RNA·DNA hybridizations are more complex for the following reason. Cellular RNA is not only a heterogeneous mixture of hundreds of species, but each contains a unique number of molecules. In *E. coli* there are at least 1000 times more molecules of the most abundant mRNA than of the most rare with a continuous distribution between these limits. The rRNA's are 50 to 100 times more frequent than the most

that such DNA contains many
f calf DNA was estimated to
00,000 times while the remain-
que sequences (*15*).
ns regarding the nature of the
be second order (*3, 12–14, 67*)
Eq. (9). Second, the reaction
sumed to be zero in Eq. (6).
mption can be made without
ld be recognized that it is not
the collision rate (*15*), so that
ed if one could only wait long
ucial for the interpretation of

' how would the unique se-
que sequences were annealed
ms of DNA were used (*15*).
otide pairs per cell than does
g calf DNA/ml used in these
entrated as are the sequences
ration should be adequate to
other extreme, the repetitive
educed thermal stability (*15*,
alf DNA per milliliter, these
concentration (equivalent to
00 copies of each sequence).
ion of complex. However, a
dissociates very quickly (sim-
t which it was formed at the
monstrates the obvious rever-
ecies. Of course, the reduced
ly a fraction (perhaps half)
e repetitious DNA segments
5).

of RNA with DNA

ns are more complex for the
a heterogeneous mixture of
nique number of molecules.
molecules of the most abun-
tinuous distribution between
more frequent than the most

common mRNA species (*43*). As a result, at any given RNA/DNA input, almost all the RNA species either are in much higher or in much lower concentration than are their complementary DNA sites. Because of this, these species should anneal with approximate first-order rather than second-order kinetics. The reaction is "pseudo" first order because the concentration of one component remains essentially constant. Of course, those species whose concentration is close to that of their complementary DNA sites will react with second-order kinetics. The net reaction kinetics will be a function of the amount of radioactivity in the two classes of RNA.

In a first-order reaction the time to complete the reaction is independent of initial concentration of the limiting component. When $b \gg a$ in Eq. (6),

$$\frac{dx}{dt} = k_1 b(a - x) - k_2 x \tag{10}$$

Integration and rearrangement gives

$$k_1 b + k_2 = \frac{1}{t} \ln \frac{Xe}{Xe - x} \tag{11}$$

with Xe the amount of a bound at equilibrium. When $x = \frac{1}{2}Xe$,

$$t_{1/2} = \frac{0.693}{k_1 b + k_2} \tag{12}$$

The time required for binding half (or any fixed fraction) of the amount of a that will ultimately be bound is a constant. The reaction will be about 97% complete after five such half-lives regardless of the initial concentration of a, the limiting component.

This relationship is very important for an evaluation of the reaction of RNA with DNA. Unfortunately, no one seems consciously to have tested this prediction directly. Many papers show a proportionality between concentration of both reactants and initial reaction rate and conclude that the reaction is second order. However, the reaction may proceed more rapidly as the concentration of the *limiting* component is increased because the final yield is increased proportionately, i.e., one is simply titrating the excess DNA sites without changing the $t_{1/2}$ of the reaction. This could explain the early observation of Nygaard and Hall (*68*) that the reaction was proportional to *both* RNA and DNA concentration. The curves (Fig. 2 of ref. *68*) showing rates as a function of DNA concentration are all initial rates of reaction with excess RNA (except 25 μg/ml DNA) since infected cell RNA starts saturating T2 DNA sites at RNA/DNA input of about one (unpublished observations). Thus, the relationship with DNA concentration probably has nothing to do with

the reaction kinetics as such, i.e., times required to complete the reactions. However, in contrast, the effects of *RNA* concentration on rate were studied using a considerable excess of most RNA label and probably did reflect changes primarily in the $t_{1/2}$ of the reaction; the RNA dependency was observed again by Bishop (*69*).

More clear cases are seen in some of the kinetic studies of stable RNA to DNA hybrid formation. Gillespie and Spiegelman (*30*) followed reactions with variable amounts of RNA and DNA; in all cases, there was an 8-fold or greater excess of stable RNA. Of course, the yield of hybrid (% of DNA complexed) was the same in all cases but the initial rates were quite proportional to RNA concentration; they were not at all related to DNA concentration or to the product of RNA and DNA concentrations. The same relationship was found for the initial rates of rRNA hybridization in ranges of $\geq$ 6-fold excess rRNA to rDNA sites (*48*). Unfortunately, I have not found any cases showing a clear rate-dependence on DNA concentration alone when it is in great excess. As indicated above, this would be difficult to show using any cell RNA with total cell DNA. It would require extremely low RNA/DNA inputs in order to give excess DNA for all species; if this condition were not met, changes in rate would be influenced by titration of the DNA, i.e., the fraction of RNA bound at equilibrium would change. DNA enriched for a specific bacterial operon would be suitable for such a study. The nearest to an excess DNA case is a DNA·DNA reaction on filters. McCarthy and McConaughy reacted increasing amounts of filter-bound DNA with a constant and limiting amount of DNA in solution (*27*). The initial rate was proportional to the concentration of filter-bound DNA, but for reasons that are not obvious, only up to about a 12-fold excess of filter-bound to free DNA.

It seems likely that the hybridization reaction, like other chemical reactions whose rate is concentration dependent, follows first-order kinetics when one nucleic acid is in considerable excess. This could be verified by standard kinetic analyses. It means that those RNA species for which complementary DNA sites are in excess all reach equilibrium at the same time regardless of how infrequent they are. Of course, they may not be fully bound at equilibrium (Section II), but if so, it would not be for kinetic reasons.

Hybridizations are often performed at an RNA/DNA ratio of about 1/10. Consider the expected kinetics for the reaction of *E. coli* nucleic acids as estimated from the following observations (*43*). At this input about 55% of the pulse-label will form hybrid of which 5% is in stable RNA and the remainder in mRNA. Assume that all mRNA species that are $\leq$ one-tenth the concentration of their DNA sites will react with first-order kinetics. This would include mRNA's that start saturating DNA

quired to complete the reactions.
VA concentration on rate were
ost RNA label and probably did
he reaction; the RNA depend-
).
he kinetic studies of stable RNA
. Spiegelman (*30*) followed re-
id DNA; in all cases, there was
. Of course, the yield of hybrid
n all cases but the initial rates
ration; they were not at all re-
duct of RNA and DNA concen-
d for the initial rates of rRNA
ess rRNA to rDNA sites (*48*).
s showing a clear rate-depend-
is in great excess. As indicated
using any cell RNA with total
RNA/DNA inputs in order to
ndition were not met, changes
f the DNA, i.e., the fraction of
e. DNA enriched for a specific
ich a study. The nearest to an
tion on filters. McCarthy and
f filter-bound DNA with a con-
ition (*27*). The initial rate was
-bound DNA, but for reasons
2-fold excess of filter-bound to

reaction, like other chemical
dependent, follows first-order
iderable excess. This could be
neans that those RNA species
n excess all reach equilibrium
uent they are. Of course, they
ection II), but if so, it would

an RNA/DNA ratio of about
he reaction of *E. coli* nucleic
servations (*43*). At this input
brid of which 5% is in stable
e that all mRNA species that
ir DNA sites will react with
VA's that start saturating DNA

sites at an RNA/DNA input of one and all less-frequent species—about 50% of all mRNA with 50% of the mRNA pulse-label (C_m/R_m in Bolton-McCarthy analysis) (*44*). Assume for simplicity that $k_1 \gg k_2$ in Eq. (10) so that all RNA is bound at equilibrium; thus for any species, $dx/dt = k_1 b(a - x)$ where b = DNA concentration, a = initial RNA concentration and x = amount of hybrid. Integration and rearrangement give

$$t = \frac{1}{k_1 b} \ln \left(\frac{a}{a - x} \right) \tag{13}$$

with

$$t_{1/2} = \frac{0.693}{k_1 b} \tag{14}$$

Thus, those species that are not fully bound at equilibrium would take somewhat longer to reach equilibrium since k_2 in Eq. (12) is opposite in sign to k_1. However, this factor should be of marginal significance for any species contributing measurable amounts of label to the hybrid. The reaction given by Eq. (13) would approach equilibrium as shown by the dashed curve of Fig. 6.

Another class of reactions would include those species that react with second-order kinetics, i.e., the RNA and DNA concentrations are within, say, a factor of 10 of each other. Assuming again that $k_1 \gg k_2$, rearrangement of Eq. (8) gives

$$t = \frac{1}{k_1} \frac{x}{a(a - x)} \tag{15}$$

A plot of the reaction with time is given by the solid curve of Fig. 6. Note that it takes much longer to reach equilibrium for reactants at the same concentration than it does when the concentration of one of the reactants is greatly reduced. The time required to anneal half of the RNA is comparable, but it takes twice as long for 75% of the RNA to react and significantly more time to reach equilibrium.

However, the rate of reaction rapidly approaches first-order kinetics as the concentrations of reactants become different. This is shown by the dot-dash curve of Fig. 6, which gives the kinetics for RNA species that are one-half the concentration of their DNA sites, i.e., in this example, they saturate at RNA/DNA inputs of 1/5 (the most abundant mRNA species of *E. coli*) (*43*). The equation generating that curve is derived from Eq. (6) (again with $k_2 = 0$)

$$t = \frac{1}{k_1(a - b)} \ln \frac{b(a - x)}{a(b - x)} \tag{16}$$

It can be seen that 98% of the RNA is bound (dot-dash curve) in less than twice the time required to bind 98% in the first-order reaction (dashed curve).

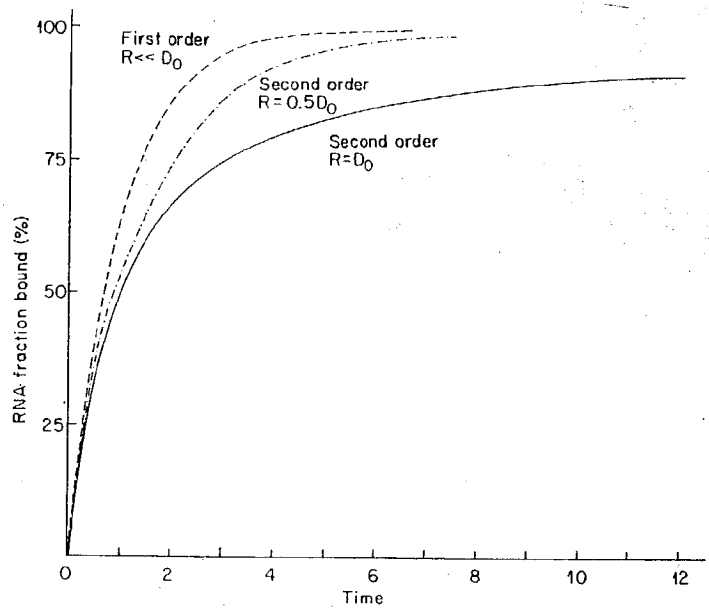
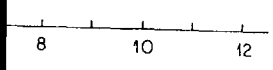


FIG. 6. The fraction of RNA bound as a function of time in first- and second-order reactions. In all cases the concentration of the DNA site is the same (D_0). In the first-order reaction, the concentration of RNA (R) is $\ll D_0$. In the second-order reactions $R = 0.5\ D_0$ or $R = D_0$.

A third class of reactions includes those RNA species whose concentrations far exceed the concentrations of their homologous DNA sites. At an RNA/DNA of 1/10, these obviously would include only the stable RNA's. These would react with the first-order kinetics of Eq. (13) except now b = RNA input and a = DNA input. Since b in this class is greater than b in the first class, these should be the fastest reacting species in the total RNA population.

At an RNA/DNA input of 1/10, the slowest reacting sequences would be those for which RNA = 0.5 (DNA). Even these would reach 90% binding almost as fast as the more rare species that react with first-order kinetics. Of course, the fraction of label that reacts with first versus second-order kinetics will vary with RNA/DNA input. There will be certain critical ratios at which a significant fraction of label will react with second-order kinetics; one such ratio would be RNA/DNA = ¼ at which the most abundant mRNA's just equal the concentration of their DNA sites (*43*). However, at almost all RNA/DNA inputs, the bulk of the radioactivity would participate in reactions in which one component is in sufficient excess to give approximate first-order kinetics.



tion of time in first- and second-
ne DNA site is the same (D_0). In
(R) is $\ll D_0$. In the second-order

RNA species whose concen-
heir homologous DNA sites.
would include only the stable
r kinetics of Eq. (13) except
ince b in this class is greater
fastest reacting species in the

est reacting sequences would
Even these would reach 90%
ies that react with first-order
hat reacts with first versus
NA input. There will be cer-
tion of label will react with
be RNA/DNA = ¼ at which
concentration of their DNA
NA inputs, the bulk of the
which one component is in
der kinetics.

If the above assumptions are correct—and they have yet to be tested rigorously but evidence so far is fairly supportive—it means that, in general, hybridization reactions should reach equilibrium much sooner than do DNA reannealing reactions for which the reactants (+ and − strands) are exactly equal in concentration. Furthermore, unlike a second-order reaction, the time to reach equilibrium is independent of the concentration of limiting component; it is not true that the very rare species would become bound if we could just wait long enough. They reach equilibrium as soon as more concentrated species but may not be fully bound because of thermodynamic, not kinetic, considerations.

Finally, the rate of the reaction is dependent on concentration of the excess component. At double the concentration the reaction will reach equilibrium about twice as fast (Eq. 13). Of course, the same applies to both reactants in a second-order reaction. This provides a means for determining if the reaction is essentially complete in a given time. It is clear that the annealing reaction of bacterial, as well as certain classes of eucaryotic cell RNA, *can* be complete by 20 hours provided the concentrations of reactants are reasonably high (*30, 31, 49, 69–75*). This was verified over a wide range of RNA/DNA inputs even with nucleic acid concentrations so low that most RNA·DNA pairs were not associated at equilibrium (*31*).

IV. Estimating Relative Transcribing Activities of DNA Sites

Most investigators use hybridization to obtain a simple "yes-or-no" answer to some question. Such questions usually ask whether some cell or tissue is making *any* RNA complementary to a given DNA. Often the investigator can ignore any chemical and technical aspects of the reaction and obtain the right answer. However, as soon as the biological question includes a quantitative answer, e.g., 57% of the RNA has some characteristic, the chemistry of the reaction becomes very important to the interpretation.

A. A Simple Case: Fraction of Genome for rRNA

Probably the most simple case is to estimate the fraction of DNA complementary to rRNA. The rRNA can be purified virtually free of other species, then hybridized at increasing RNA/DNA inputs to determine the input at which all complementary sites are just filled. From this value and the fraction of RNA hybridized, it is a simple calculation to derive the fraction of DNA complementary to rRNA (*24, 43, 71, 73, 74, 76, 77*).

B. A Complex Case: Total Cell RNA

It is much more difficult to estimate the relative numbers of molecules per species in total RNA. At a given RNA/DNA input, a certain fraction of the RNA will anneal to all copies of complementary DNA and be in excess. McCarthy and Bolton (*44*) derived an ingenious procedure for estimating the fraction of input DNA complementary to this fraction of RNA; a plot of these fractions gives the relative amount of RNA complementary to a unit of DNA for any fraction of the genome. The analysis can be restricted to the mRNA (*43*), a minor component of the RNA mass, which then increases the resolution immensely since the bulk of the gene sites of bacteria or phage, at least, code for mRNA (Fig. 7).

An important limitation of such an analysis is that one can derive the fraction of the DNA that codes for a certain fraction of mRNA only if the fraction of radioactivity in the mRNA species corresponds to the frac-

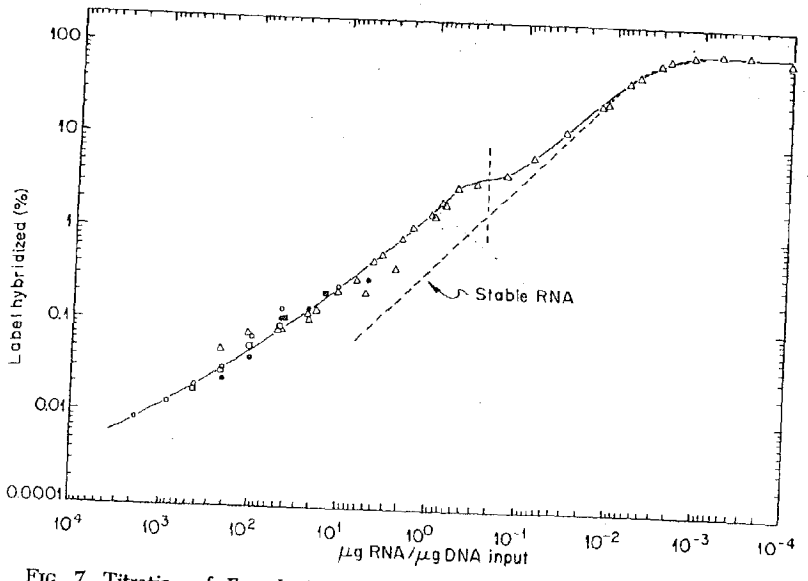
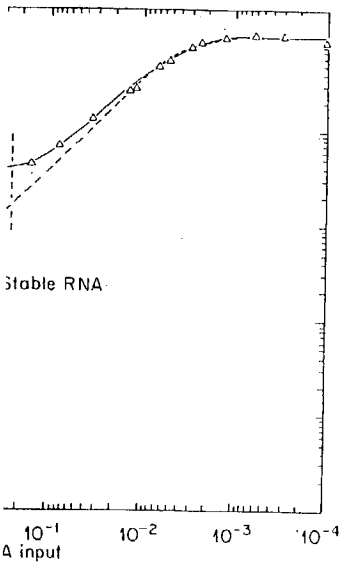


FIG. 7. Titration of *E. coli* denatured DNA with total, or long-labeled, *E. coli* RNA. The percentage of input label that forms stable hybrid is plotted against RNA/DNA input. The maximum is set at 100%. For comparison, the titration of the stable RNA is reproduced as the dashed line. The difference between these curves gives the fraction of input RNA that is hybridized unstable RNA. Note the effect of nonspecific sticking when the hybridizations occur on filters at RNA/DNA inputs > about 50. (△) Hybridization with DNA filters containing a known amount of DNA (2, 5, 15, 50, or 200 μg). Hybridization with 2 (○), 5 (□), or 10 μg (●) DNA in liquid (reference 43).

e relative numbers of molecules
/DNA input, a certain fraction
omplementary DNA and be in
ed an ingenious procedure for
nplementary to this fraction of
relative amount of RNA com-
on of the genome. The analysis
minor component of the RNA
mmensely since the bulk of the
ide for mRNA (Fig. 7).
lysis is that one can derive the
ain fraction of mRNA only if
pecies corresponds to the frac-



vith total, or long-labeled, *E. coli*
stable hybrid is plotted against
or comparison, the titration of the
e difference between these curves
unstable RNA. Note the effect of
on filters at RNA/DNA inputs >
taining a known amount of DNA
○), 5 (□), or 10 μg (●) DNA

tion of mRNA mass in this species. This condition is only fulfilled in steady-state labeling, i.e., long-labeled RNA for which all molecules have the same specific activity. It is easily attained when working with microorganisms but is difficult to achieve in many eucaryotic cell studies.

The titration of DNA by pulse-labeled RNA can be of interest if it is compared to the titration by long-labeled RNA. If the titration curves are the same, i.e., a given fraction of DNA is complementary to the same fractions of pulse-label and RNA mass, it suggests that on the average the number of molecules per species is determined by rate of synthesis rather than stability. This was concluded with respect to mRNA of *E. coli* (*43*). We may note that this conclusion is supported by recent observations that the rate of chemical degradation of one *E. coli* mRNA (*lac* mRNA) is identical to the average rate of degradation of the total mRNA (*50*).

An alternative way to analyze the data from a titration is to use the equations of a Langmuir adsorption isotherm for analysis of the filling of a surface. In this case, the surface is DNA, and it is being filled by RNA. The equilibrium constant

$$K = \frac{(R - S)(aD - S)}{S}$$

where R = RNA input, D = DNA input, S = RNA bound, and a = the fraction of DNA competent to accept RNA. Rearrangement gives

$$D/S = \frac{k}{a(R - S)} + \frac{1}{a}$$

A plot of D/S against $1/R - S$ has a slope of K/a and an intercept of $1/a$. The data from the titration of the *E. coli* chromosome (*43*) were replotted this way, and as expected, the slope changed as a function of RNA/DNA input. This follows from the fact that the reaction is the summation of hundreds of sites being filled, and as a result, a is really not a constant. However, the intercept of $1/a$, extrapolated to correspond to infinite RNA input, gave the same value for a as had been derived by the McCarthy-Bolton analysis: 10% of the DNA. This value of 10% was interpreted to mean that all the mRNA detected over the range of abundances measured (about 2000-fold) was transcribed from 20% of the potential gene sites (*43*).

V. Quantitative Limitations of Hybridization

Recently, Bishop *et al.* challenged the use of hybridization for quantitative studies (*48, 69, 70, 78*). Basically, two considerations seem to be under question. The first is a kinetic one: the reaction is not complete in

20–24 hours because as "reacting sequences are withdrawn into DNA·RNA hybrid, their concentration falls and the reaction rate falls." The second is simply a flat statement that the amount of RNA hybridized does not bear a proportional relation to RNA input (*70*).

A. Kinetic Considerations

The first criticism is discussed in Section III, B. As pointed out there, it is true that, at a given concentration and ratio of reactants, a certain fraction of labeled RNA that will ultimately be bound may still be free at 20 hours. However, the reaction rate will be a function of the concentrations of either one or both reactants depending upon whether the reaction rate is first or second order. The effects of nucleic acid concentrations on reaction rates are well known (*30, 68–70*). However, it *follows* that at a given RNA/DNA input the reaction of bacterial and certain eucaryotic nucleic acids *can* reach equilibrium by 20 hours if the nucleic acid concentrations are reasonably high. This has been observed to be the case in many studies (*30, 31, 49, 69–75*).

There is another way to show that virtually all radioactive RNA for which complementary sites are available has annealed within the reaction time. This can be concluded if there is no increase in the fraction bound when the concentration of reactants is increased at a given RNA/DNA input. This again follows from the fact that the rate is a function of the concentration of either one or both nucleic acids. Thus, the yield after a limiting time should be increased by increasing concentrations if such a yield is limited for kinetic reasons. For example, concentrations giving "plateau" values for fraction bound were used in the titration of the *E. coli* chromosome (Fig. 1).

B. Thermodynamic Considerations

The second criticism states that the amount of RNA bound is not proportional to added RNA. It is based mainly on results of hybridizations with partially defined mRNA such as that from the galactose (*70*) or tryptophan (*61, 79*) operons and with tyrosine tRNA (*80, 81*). It is to be noted that there is no sharp transition in the curve of amount of hybrid versus RNA input at that RNA/DNA input at which the DNA sites are just saturated.

With respect to the mRNA studies, there are at least two possible explanations for this result. First, a given species of bacterial mRNA is undoubtedly heterogeneous with respect to size. It has been shown for *trp* mRNA (*82*) and for *lac* mRNA (*50*) that degradation can commence before synthesis terminates. Degradation occurs on a random basis so that some molecules may never exceed triplet nucleotide length while others exist at the maximum size for several minutes (*50*). Thus, at the

ces are withdrawn into DNA·
d the reaction rate falls." The
mount of RNA hybridized does
nput (*70*).

ın III, B. As pointed out there,
ıd ratio of reactants, a certain
ely be bound may still be free
ll be a function of the concen-
pending upon whether the re-
ects of nucleic acid concentra-
9, 68–70). However, it *follows*
ction of bacterial and certain
ium by 20 hours if the nucleic
ıis has been observed to be the

tually all radioactive RNA for
ıs annealed within the reaction
increase in the fraction bound
reased at a given RNA/DNA
ıt the rate is a function of the
: acids. Thus, the yield after a
asing concentrations if such a
xample, concentrations giving
ısed in the titration of the *E.*

ıunt of RNA bound is not pro-
y on results of hybridizations
t from the galactose (*70*) or
ine tRNA (*80, 81*). It is to be
ıe curve of amount of hybrid
t at which the DNA sites are

ere are at least two possible
species of bacterial mRNA is
ı size. It has been shown for
at degradation can commence
ıccurs on a random basis so
iplet nucleotide length while
ıl minutes (*50*). Thus, at the

RNA/DNA inputs at which DNA sites are just being filled, sites that are partially filled with incomplete lengths of mRNA may be further filled by another small mRNA that is nonoverlapping with the first. The net result of such a terminal filling would be to make the transition from excess to limiting DNA less sharp.

A second and more important consideration applies to such studies with any defined RNA. As discussed in Section II, at any given RNA/DNA input the yield of hybrid at equilibrium is a function of concentration of reactants. Thus, at sufficiently low concentrations of reactants, not all the RNA molecules are bound for which complementary DNA sites are available (Fig. 1). In the gal mRNA·DNA work (*70*), it is not possible to evaluate the effect of concentration since the concentrations are not given. However, in the saturation of tyrosine tRNA sites, which is considered a very clear case (*70, 78*), the reactant concentrations are given (*80, 81*). In the Landy *et al.* study (*80*), 2 μg of ϕ80dSu3 + DNA is saturated by 0.8 μg of RNA input. At this ratio, 0.0016 μg of tRNA is bound to the same amount of DNA (0.08%). These are extremely low concentrations of reactants. The concentrations in the other study are higher, but still very low (*81*). We observed obvious concentration dependence on yield with 0.06 μg of rRNA reacting to 0.56 μg of rDNA (*43*) and with about 0.001 μg of *lac* RNA to 0.2 μg of *lac* DNA (Fig. 1). In fact, the only obvious interpretation of a curvilinear approach to saturation is that there is a concentration dependence for yield. It means that more RNA has to be added to fill DNA sites that are available. As emphasized earlier (*31* and Section II), the best way to prove that this dependency exists is to choose *one* RNA/DNA input and measure fractional yield as a function of input concentrations. In contrast, by adding more RNA *per* DNA one may also be observing the actual titration of sites as well as a changing "efficiency" of reaction.

In fact, some saturation curves have shown remarkably sharp transitions in slope when going from unsaturated to saturated DNA sites, e.g., the saturation of rRNA sites (*71, 83* and unpublished observations) and even the saturation of *trp* mRNA sites (if the curve of Fig. 5, reference *61*, is redrawn to fit the points better).

A further argument for the inadequacy of hybridization for quantitative estimations is based on a complete titration of the *E. coli* chromosome. Instead of concluding that only 20% of the potential gene sites make significant levels of RNA (*43*), the author arrived at a value of 28%, and furthermore felt that even this figure might be low (*70*). A very important technical consideration throws doubt on these results. With 2 mg/ml RNA input, 0.12 μg of RNA was hybridized per microgram of DNA. Unfortunately, the actual amounts of RNA and DNA reacting are not given, but if it were 1 μg of DNA, this would give 0.006% of the RNA

hybridized; if 10 μg of DNA, 0.06% hybridized. It seems unlikely that more than 10 μg of DNA was used, since this would give an RNA/DNA input of 200, which is still far from saturating all DNA sites (*43*). As pointed out in the earlier study (*31*, *43*), the problem of nonspecific adsorption becomes extreme at these very high RNA/DNA inputs *especially* with the long-labeled RNA and the filter method used in these experiments (*70*). At RNA/DNA inputs greater than about 20, reaction in solution, was used to give about a 10-fold lower blank (no DNA) (*31*). However, even with this procedure, it was not possible to observe complete saturation of DNA by long-labeled RNA (Fig. 7) (pulse-labeled RNA for which blank adsorption is less significant did approach saturation) (*43*). The no-DNA controls are not given but in another paper (*69*) they are stated to be 0.05%—a very respectable value. However, while respectable, I doubt if they are so reproducible as to warrant sweeping generalizations from values that are barely above this background.

As stated in the earlier study (*31*), the hybridization reaction appears to obey the law of mass action (*84*). Second, independent of this thermodynamic statement, there is the kinetic consideration: the rate of the reaction will be a function of either one or both reactant concentrations depending upon whether it is first or second order. Certain kinetic considerations have been recognized since the reactions were first performed. However, the first consideration has been ignored or avoided even though it can explain an enormous number of apparently contradictory results; it becomes even more obvious with eucaryotic cell work, where the degree of reversibility of the reaction is measured by the variable stabilities of hybrids formed with variable degrees of mismatching (Section VII). Note that the kinetic and thermodynamic variables can be completely separated experimentally, e.g., a reaction can have reached equilibrium without having all RNA hybridized for which DNA sites are available. Fortunately, since both the kinetic and thermodynamic limitations are a function of reactant concentrations, the maximal yield of hybrid can be determined for any RNA/DNA input by determining those concentrations that give the maximal yield in a given time. Rather than reject use of hybridization for quantitative work because of its complexity, it is important to recognize these and other limitations and adjust reaction conditions accordingly.

VI. Use of Competitor RNA to Estimate Specificity

Two identical molecules will compete with each other for a common binding site, if that site is limiting. This simple principle has been used very frequently to assess similarities between two RNA populations com-

ed. It seems unlikely that more
ould give an RNA/DNA input
ll DNA sites (*43*). As pointed
blem of nonspecific adsorption
A/DNA inputs *especially* with
iod used in these experiments
about 20, reaction in solution,
nk (no DNA) (*31*). However,
ıle to observe complete satura-
(pulse-labeled RNA for which
proach saturation) (*43*). The
her paper (*69*) they are stated
Iowever, while respectable, I
rrant sweeping generalizations
background.
hybridization reaction appears
d, independent of this thermo-
ısideration: the rate of the re-
both reactant concentrations
nd order. Certain kinetic con-
reactions were first performed.
nored or avoided even though
ırently contradictory results; it
c cell work, where the degree
l by the variable stabilities of
mismatching (Section VII).
variables can be completely
an have reached equilibrium
ıich DNA sites are available.
rmodynamic limitations are a
ıximal yield of hybrid can be
ermining those concentrations
ne. Rather than reject use of
e of its complexity, it is im-
tions and adjust reaction con-

Estimate Specificity

ith each other for a common
nple principle has been used
n two RNA populations com-

peting for a common DNA. Competition by nonidentical RNA molecules will be discussed in the section on hybridization of eucaryotic cell nucleic acids.

A. Some Basic Considerations

Certain basic principles should be considered in such studies.

1. RNA molecules compete on a mass basis while usually the measurements are of radioactivity bound. Thus, if the labeled RNA molecules have different specific activities, the amount of label excluded by the competitor may bear little relation of the amount of RNA excluded.

2. Only labeled RNA molecules that are in excess of their DNA sites (without competitor) will have their binding decreased to an extent proportional to the amount of the competitor RNA molecules added.

B. A Not So Simple Case: Estimating the Fraction of Pulse-Labeled RNA That Is rRNA

Probably the simplest example of competition is directed to show what fraction of pulse-labeled RNA is ribosomal (rRNA). Ideally, one would want three different RNA samples: (a) RNA with label only in rRNA; (b) purified unlabeled rRNA; and (c) the pulse-labeled RNA ([^{3}H]RNA) being studied. The labeled rRNA should be titrated to excess DNA to determine the RNA/DNA input at which the sites for rRNA are just filled. For *E. coli* this ratio is about 1/160 to 1/200 (*43*, *49*). Thus, at RNA/DNA inputs $<1/200$, DNA sites will be in excess and proportional competition will not be observed. However, it would not be wise to use higher [^{3}H]RNA/DNA inputs, since at higher RNA/DNA input the fraction of hybrid that is mRNA is greater. For example, at RNA/DNA = 1/10, only 4% of *E. coli* pulse-labeled ^{3}H in the final hybrid is rRNA, while at 1/200 it is 50% (*43*); in the former case only 4% of the label could be excluded by excess rRNA.

How much competitor RNA should be added? Offhand, one might want to add as much as possible. This would be satisfactory if one could be sure that the purified rRNA competitor is not contaminated by other RNA species. For example, if the competitor RNA/[^{3}H]RNA were 1000 and 0.2% of the competitor were mRNA, then 2 units of [^{1}H]mRNA would be present per 200 units of DNA when the [^{3}H]RNA/DNA = 1/200. This would be sufficient to saturate *E. coli* sites for a significant fraction of the most abundant mRNA. The result would suggest that rRNA is a much larger fraction of pulse-labeled RNA than it actually is. For this reason, it is usually safer to choose a more conservative ratio of [^{1}H] to [^{3}H]RNA (say 50). Because of this, it is best to avoid an extremely low [^{3}H]RNA/DNA input (e.g., $<1/1000$) since the reduction in ^{3}H bound would not

be proportional to competitor RNA added and an additional calculation would be necessary. Thus, in this case the [^{3}H]RNA/DNA would best be close to, but not greater than, 1/200.

C. A Very Complex Case: Estimating Relatedness of RNA Populations from Different Cells

The most frequently used competitor experiments involve the competition of one total cell RNA sample against another. In this case the net competition is the summation of the simultaneous competitions of hundreds or thousands of species. This procedure has been used to study transcriptional patterns as a function of infection time by virulent phages (*85–87*) and similarities between RNA from different eucaryotic cells or tissues (e.g., *75*, *88–92*). The same general rules apply to this more complex situation, i.e., the [^{3}H]RNA/DNA input as well as the [^{3}H]RNA/[^{1}H]RNA inputs must be chosen to provide answers to a specific question.

There are several hundred species of mRNA in a cell, and probably the numbers of molecules per species also vary over a several hundredfold range. Therefore, in practice, since the competitor competes against the ^{3}H that would *otherwise* form hybrid, its effectiveness as a competitor depends *only* on its ability to compete against those [^{3}H]RNA species that form the bulk of the hybrid. The particular set of species forming the bulk of hybrid will be different at each different [^{3}H]RNA/DNA input and include more species as the [^{3}H]RNA/DNA ratio is raised.

In the limiting case at very high [^{3}H]RNA/DNA input, all DNA sites will be filled, and all RNA species competed against. In theory this would be the condition from which to make generalizations about the relatedness of the active DNA sequences. Each DNA site would be filled with one [^{3}H]RNA molecule (no competitor) irrespective of the frequency distribution of RNA molecules per species. In practice this measurement is difficult for technical reasons; enormous amounts of RNA are required and the fraction of ^{3}H hybridized becomes intolerably low.

Some competition experiments have tended to the opposite extreme, i.e., very low [^{3}H]RNA/DNA inputs (say $<1/10$) to estimate "similarities" between total RNA from cells or tissues. A few comments should be made. First, ribosomal and transfer RNA comprise the bulk of the RNA in all cells, while the total mass of mRNA is a minor fraction (about 1% to 8%). Thus, the *E. coli* case (mRNA = 3% of RNA), should serve as a satisfactory example for most cells. At a [^{3}H]RNA/DNA ratio of 1/25, approximately 17% of the total *E. coli* RNA mass can be bound to DNA sites. Of this final hybrid, 82% would be stable RNA and only 18% would be mRNA (Fig. 7). Since RNA molecules compete on a mass basis, a competing cell RNA could eliminate over 80% of the RNA *mass* from the

l and an additional calculation
ıe [³H]RNA/DNA would best

elatedness of RNA

experiments involve the com-
ıst another. In this case the net
ltaneous competitions of hun-
:dure has been used to study
ection time by virulent phages
m different eucaryotic cells or
rules apply to this more com-
put as well as the [³H]RNA/
answers to a specific question.
ıRNA in a cell, and probably
vary over a several hundred-
: competitor competes against
s effectiveness as a competitor
;ainst those [³H]RNA species
ticular set of species forming
ıch different [³H]RNA/DNA
H]RNA/DNA ratio is raised.
VA/DNA input, all DNA sites
l against. In theory this would
lizations about the relatedness
site would be filled with one
ıective of the frequency dis-
practice this measurement is
nounts of RNA are required
intolerably low.
ded to the opposite extreme,
<1/10) to estimate "similar-
ues. A few comments should
IA comprise the bulk of the
VA is a minor fraction (about
3% of RNA), should serve as
³H]RNA/DNA ratio of 1/25,
mass can be bound to DNA
)le RNA and only 18% would
compete on a mass basis, a
% of the RNA *mass* from the

hybrid without having a single mRNA species in common. Of course, the percent of ³H eliminated would also be a function of the distribution of ³H in the stable and mRNA fractions. If 50% of the ³H were in mRNA then elimination of stable RNA from the hybrid would only exclude 14% of the ³H. However, in many eucaryotic cells mRNA may represent a much lower fraction of synthesized RNA (*93*). If it were 10% rather than 50%, competing out only stable RNA could exclude 46% of the ³H.

Second, again using *E. coli* as an example, regardless of the ³H distribution, there would have to be at least a 10-fold increase in the mass of the most abundant mRNA's before even they would saturate their DNA sites (Fig. 7). Such molecules might be present in the competing RNA, but if they were present at only one-tenth their concentration in the [³H]RNA, a 100-fold excess of total RNA ($^1H/^3H = 100$) would be necessary in order to saturate these sites. Clearly, such a starting [³H]RNA/DNA input might give very misleading results regarding similarities between molecules in the mRNA fraction.

Most competition experiments start with some intermediate [³H]RNA/DNA input and attempt to estimate relatedness of mRNA from a different cell or tissue. At an [³H]RNA/DNA input of 1/4, 4% of *E. coli* RNA forms hybrid of which about 37% is stable RNA. Thus, even at this ratio, 37% of the hybridized RNA *mass* could be eliminated by competitor RNA lacking mRNA in common. However, if 50% of the ³H were in mRNA, elimination of rRNA would only eliminate 0.75% of the ³H from hybrid. In this case, successful competition would indicate similarities in the mRNA. However, the bulk of the mRNA is transcribed from a small fraction of the active DNA; in *E. coli*, 80% of the mRNA is transcribed by only 10% of the active genes (*43*). As a result, at an [³H]RNA/DNA input of 1/4, 80% of the hybrid could be eliminated even though 90% of the *species* of mRNA were nonidentical. While such a result would suggest that 80% of the ³H molecules have competitors in the competitor RNA, it says very little about similarities between the species or kinds of molecules in the two samples. This simple but important distinction is usually overlooked.

Another consideration in this kind of experiment makes questionable the conclusions almost always derived from competitions of this kind. Invariably, the ratio of competing RNA to [³H]RNA is increased to some extremely high value (≥ 100). If one follows the decrease in ³H hybridized as this ratio is increased, it is clear that it does not decrease proportionately. Two variables contribute to this. First, the DNA sites for the abundant species, which contribute the bulk of the ³H hybrid, are not all filled to the same extent at the [³H]RNA/DNA input. The other reason is that the ³H counterparts of these abundant [³H]RNA may be a *minor*

fraction of the mRNA in the competitor, i.e., the frequency distributions of the same molecules may be very different in the two RNA's. Unfortunately, as a consequence of the first variable, it is almost impossible to quantitate the second. A simple example will clarify this important point.

Suppose 80% of the [^{3}H]mRNA is coded by 5% of the active genes and that identical genes comprise 5% of the active genes in the cells producing the competitor RNA. However, in these latter cells this RNA represents only 5% of the mRNA product (rather than 80%). If *no* other RNA were common to the labeled and unlabeled cells, about 70% of the [^{3}H]RNA would be excluded by competitor at an [^{1}H]/[^{3}H]RNA input = 100 (Fig. 8). The usual conclusion from such results would be that 70–80% of the RNA molecules are common to each cell. Instead, only 5% of the transcribed DNA and 5% of the mRNA molecules would actually be common to both cells. What would appear to be closely related cells or tissues would really be similar to only a minor extent either with respect to genetic sequences or transcription products.

Furthermore, almost identical competition curves are obtained from competitor RNA's with significantly different degrees of identity with the [^{3}H]RNA. The three similar competition curves in Fig. 8 describe cases in which 80% of the [^{3}H]RNA molecules have unlabeled counterparts in each of the competing RNA. For simplicity these ^{3}H molecules are in only one or two possible frequency classes and their ^{1}H counterparts are in two frequency classes. The less frequent molecules saturate their DNA sites at higher RNA/DNA than do the more frequent. The concentration of any class is determined by *both* the ^{3}H and the ^{1}H molecules (the two variables alluded to above). As the number of frequency classes increases for both the ^{3}H and the ^{1}H populations, so do the possible number of distributions that could give almost the same curve. This is the probable situation with RNA from any two different cells. As can be seen from the oversimplified cases in Fig. 8, competitor RNA with only 5% of its molecules identical to molecules in the labeled RNA can give a curve with the same slope and end point as one with 30%.

Clearly, the results of competition experiments will depend to a major extent on the initial choice of [^{3}H]RNA/DNA input. While this widely used procedure has been of some value to assess relatedness, results of still greater value would be obtained if it were used in a more analytical fashion.

Finally, consider a typical competition experiment in which high concentrations of cell RNA-2 exclude, say, 50% of the ^{3}H from cell RNA-1 from hybrid formation. The incorrect interpretations that are commonly

i.e., the frequency distributions
rent in the two RNA's. Unfor-
uriable, it is almost impossible
ple will clarify this important

l by 5% of the active genes and
tive genes in the cells produc-
ese latter cells this RNA rep-
ıer than 80%). If *no* other RNA
ɔeled cells, about 70% of the
itor at an [^{1}H]/[^{3}H]RNA in-
ı from such results would be
mon to each cell. Instead, only
RNA molecules would actually
ɔear to be closely related cells
a minor extent either with re-
n products.
ion curves are obtained from
rent degrees of identity with
ion curves in Fig. 8 describe
ɔs have unlabeled counterparts
city these ^{3}H molecules are in
and their ^{1}H counterparts are
molecules saturate their DNA
'e frequent. The concentration
nd the ^{1}H molecules (the two
of frequency classes increases
so do the possible number of
ıe curve. This is the probable
ıt cells. As can be seen from
itor RNA with only 5% of its
ed RNA can give a curve with
ı 30%.
iments will depend to a major
)NA input. While this widely
assess relatedness, results of
'ere used in a more analytical

xperiment in which high con-
ℓ of the ^{3}H from cell RNA-1
ɔretations that are commonly

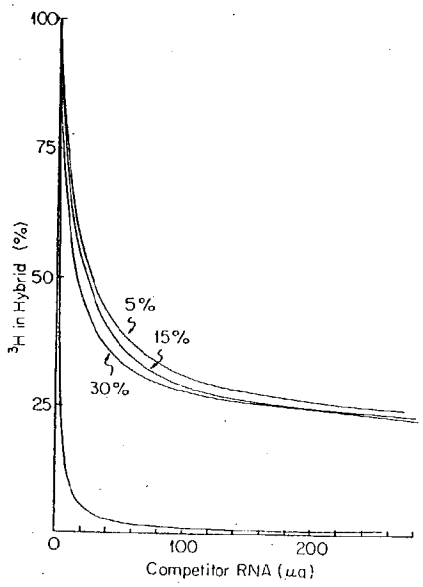


FIG. 8. The percentage of [^{3}H]RNA hybrid as a function of competitor [^{1}H]RNA concentration (100% = value with no competitor). One microgram of [^{3}H]RNA is hybridized to DNA from which it was transcribed. The lower curve represents competition by the homologous unlabeled RNA. In the other three curves, 80% of the [^{3}H]RNA is in species present in the [^{1}H]RNA. (5%) of the competitor [^{1}H]RNA competes with 80% of the [^{3}H]RNA for sites that are all saturated with no [^{1}H]RNA present. (15%) of the competitor RNA competes for DNA sites for 80% of the [^{3}H]RNA. There are two abundance classes for these molecules, and the ^{3}H molecules all have the same specific activities. The first class contains 70% of the ^{3}H, and the second 10%; 5% of the [^{1}H]RNA is homologous to the first ^{3}H class and 10% to the second. The [^{3}H]RNA/DNA is such that the first class would just saturate its DNA sites. (30%) of the competitor RNA is homologous to 80% of the [^{3}H]RNA. There are two frequency classes as in the previous case, except that the 70% and 10% classes of ^{3}H are now 60% and 20%, respectively, and the 5% and 10% ^{1}H classes are now 10% and 20%, respectively.

made are: (a) 50% of the RNA molecules being made are common to both cells. This is unjustified for reasons given above. (b) 50% of the *species* of RNA are identical. This interpretation is even further from reality than the first but invariably authors casually substitute species distributions for molecule distributions at some point. It is incorrect because most of the label being excluded from hybrid is in a small fraction of the total species. (c) Finally, more cautious investigators add the phrase "on the average" to cover the possibility that the 50% of the ^{3}H

excluded could be in some group of RNA with widely different numbers per species both for the labeled RNA and unlabeled competitor. While this phrase probably does protect the conclusion because it is undefined (average of what?), it takes all significance out of the result.

An analogy comes to mind. The *average* year round temperature in St. Louis is listed as a comfortable 59°F. This fact is unimpeachable; it is <20°F most of four months, >90°F most of another four, and everything in between during the remainder.

VII. Hybridization of Nucleic Acids from Eucaryotic Cells

A. Factors Adding to the Complexity of Interpretations

The importance of the hybridization reaction is most severely tested in the reactions of RNA and DNA from eucaryotic cells. In general the more complex an organism, the more DNA per cell (haploid set) of the organism (*15*). The haploid DNA content of a mammalian cell may be 1000 times greater than that of *E. coli*. The work of the Carnegie group, especially, has shown that a fraction of the eucaryotic RNA·DNA complex formed during most hybridizations has a relatively low thermal stability (*46, 47*). It is this fraction that includes DNA sequences that reanneal much more rapidly than expected from kinetic considerations (*15, 27*) (Section III). Both these observations led to the proposal that a fraction of eucaryote DNA contains stretches of nucleotides that are repeated hundreds or thousands of times and probably differ from each other to a relatively minor extent (*15, 27, 45, 46, 94*). Britten and Davidson proposed a model that assigns a function to the repetitive sequences; the model takes the original Jacob-Monod operon system (*95*) and amplifies it to include multiply interacting operons (*96*). Much of the RNA that contains repetitive sequences may include the rapidly turning-over RNA that is confined to the nucleus; this nuclear RNA seems to represent the bulk of RNA being synthesized in the cell (*93, 97–101*). Such RNA could play a regulatory role by binding to specific operator sites on the chromosomes (*102, 103*).

The presence of repetitive sequences increases the complexity of the reactions of DNA with DNA or of DNA with RNA. However, it does not make these reactions of any less value for the study of gene expression in eucaryotic cells than they are for such studies with bacteria and phage. The same chemical reaction is involved, and the same general considerations regarding rates and yields should still apply.

with widely different numbers
d unlabeled competitor. While
clusion because it is undefined
ce out of the result.
'ge year round temperature in
This fact is unimpeachable; it
ost of another four, and every-

s from Eucaryotic Cells

f Interpretations

action is most severely tested
ucaryotic cells. In general the
per cell (haploid set) of the
of a mammalian cell may be
work of the Carnegie group,
eucaryotic RNA·DNA com-
has a relatively low thermal
ncludes DNA sequences that
d from kinetic considerations
ions led to the proposal that a
es of nucleotides that are re-
id probably differ from each
5, *46*, *94*). Britten and David-
n to the repetitive sequences;
operon system (*95*) and am-
rons (*96*). Much of the RNA
ude the rapidly turning-over
clear RNA seems to represent
cell (*93*, *97–101*). Such RNA
specific operator sites on the

creases the complexity of the
h RNA. However, it does not
e study of gene expression in
ies with bacteria and phage.
the same general considera-
apply.

B. Stringency of Reaction Conditions and Thermal Stabilities of Complex

Partially mismatched hybrids are readily formed when most, but not all, of the nucleotides of RNA molecules are complementary to certain DNA strands. The T_m may be reduced about 1° for every 1% increase in mismatching of the complementary strands (*26*, *104*). However, it should be emphasized that the extent to which mismatched hybrids can form is relative and dependent upon the "stringency" of the reaction conditions. The stringency refers to the extent to which the reaction conditions allow only completely complementary structures to form. Generally, stringency is proportional to temperature and inversely proportional to salt concentration.

Denis (*105*) showed that amphibian hybrids are of variable stability, and Church and McCarthy (*45*) compared the reactions of viral, bacterial, and mammalian nucleic acids. *B. subtilis* DNA appears to contain no repeating sequences (*27*); however, mismatched hybrids are formed when the reaction occurs at 50° rather than $\geq$60° (*45*). Partial mismatching was detected by an examination of the thermal profile for elution of the RNA from the hybrid. In general, the less stable a hybrid, the lower is its T_m for elution. It follows that the more variable the stability of total hybrid, the broader is the temperature range of elution. As the temperature of hybridization is raised, hybrid stability and homogeneity with respect to stability both increase when mouse RNA and DNA react. At 75° the melting curve was as steep as the steepest *B. subtilis* curve and had a T_m of about 80° in 0.15 M NaCl + 0.015 M Na citrate (*45*).

Analyses of thermal elution profiles are therefore of extreme importance, especially in hybridization studies with eucaryotic cells. In fact, in most eucaryote cell hybridizations, they are absolutely essential in order to establish the significance of the hybrid yield. Clearly, 20% of the RNA hybridized with variable degrees of mismatching means something different than 5% of the RNA all hybridized with complete complementarity. Both results could be obtained at the same RNA/DNA input with different conditions of salt and temperature. Similarly, estimates of the extents of similarity of two RNA populations by competition experiments are greatly affected by reaction conditions (*45*). This follows directly from the preceding considerations: an RNA molecule that is similar, but not an exact copy of another RNA, will compete for a common DNA site if the stringency conditions are sufficiently relaxed.

Those hybrids with lower association constants due to mismatching should elute from the DNA at lower temperatures. In theory, both the

shape of the elution curve and the average T_m should be changed if a significant fraction of labeled hybrid is in species whose concentration is sufficiently low to show a reduced T_m due to concentration. This might never be noticed since enormous amounts of DNA need to be reacted to anneal the unique species. However, once formed in solution, a small fraction of hybrid trapped on a filter could show this concentration effect.

C. Expected Reaction Kinetics for Hybridization of Mouse RNA to DNA

The reaction kinetics for eucaryotic cell hybridizations deserves special comment. The same basic considerations outlined in Section III should apply. Assume a reaction of mouse RNA to mouse DNA. It was estimated that about 10% of mouse DNA is composed of partially repeating sequences ("satellite" DNA) repeated about 1,000,000 times. At the other extreme, 70% of the DNA is composed of unique sequences, and the remaining 20% of sequences are repeated from two to 1,000,000 times each (*15*). The kinetics and yields of hybrid will depend not only on the concentrations of DNA sites, but on the concentrations of RNA species for these sites as well. The fraction of the genome that is transcribed in any given time is probably very small (*93, 98*). Let us assume it is 1% in this case and further that each of the above gene classes synthesizes RNA at the same total rate, i.e., the satellite DNA accounts for 10% of the RNA made and the other two classes for 70% and 20%, respectively.

Consider the satellite DNA first. Ten percent of the pulse-labeled RNA is made by 0.1% of the total DNA. However, if it is true that a large part of the repetitive DNA is for rapidly turning-over nuclear RNA having a regulatory function (*93, 96*), then this RNA represents a much smaller part of the RNA mass (e.g., 0.1%). Furthermore, if the reaction conditions are not sufficiently stringent, this 0.1% of the RNA mass (with 10% of the pulse label) will form hybrids with varying degrees of stability with 10% of the DNA even though it had been derived from only 0.1% of the DNA. Therefore, at an RNA/DNA input of one, there should be a great excess of DNA for even the most abundant RNA species in this class.

For this reason these molecules will react with pseudo-first-order kinetics. The time required for binding half the RNA (or any fraction) will be approximately inversely proportional to DNA concentration. The fraction of this RNA bound will be a function of the stringency of the reaction conditions. The time to complete the reaction should be independent of RNA concentration so that even the rarest RNA molecules will be fully bound as soon as the most abundant.

The reaction of the RNA to the class of unique DNA sequences will follow different kinetics. These molecules contain 70% of the pulse label

ge T_m should be changed if a
species whose concentration is
e to concentration. This might
of DNA need to be reacted to
e formed in solution, a small
show this concentration effect.

ridization of

hybridizations deserves special
outlined in Section III should
mouse DNA. It was estimated
sed of partially repeating se-
1,000,000 times. At the other
unique sequences, and the re-
m two to 1,000,000 times each
l depend not only on the con-
entrations of RNA species for
ome that is transcribed in any
. Let us assume it is 1% in this
ne classes synthesizes RNA at
accounts for 10% of the RNA
id 20%, respectively.
percent of the pulse-labeled
wever, if it is true that a large
ning-over nuclear RNA having
NA represents a much smaller
ore, if the reaction conditions
e RNA mass (with 10% of the
ing degrees of stability with
derived from only 0.1% of the
f one, there should be a great
t RNA species in this class.
react with pseudo-first-order
f the RNA (or any fraction)
l to DNA concentration. The
tion of the stringency of the
the reaction should be inde-
en the rarest RNA molecules
dant.
unique DNA sequences will
ontain 70% of the pulse label

and are derived from 0.7% of the DNA. There is no information on the stability of this fraction, but it probably includes most of the mRNA, and in mammals there is some suggestive evidence that mRNA's are quite stable (*93*). In any case, for purposes of this example, assume the specific activity of this fraction to be one-tenth that of the preceding repetitive fraction. This would mean that the unique fraction was 7% of the RNA mass. It could be somewhat more if the tRNA were derived from unique sequences, but it is assumed that rRNA (comprising about 80% of the bulk RNA) would be derived from the intermediate class containing the fairly repetitive sequences (*24, 72–74*).

At an RNA/DNA input of one, 7% of the RNA will be reacting with 0.7% of the DNA (the unique sequences). Thus, the *average* concentration of an RNA species will be 10 times that of its DNA site. However, the concentrations of each species will vary over some wide range. Assume 1000 different species [about the number required to handle metabolism (*106*)] with a continuous distribution of numbers of molecules per species; the most abundant contain 100 times as many molecules as the most rare and account for 1% of the mass of this fraction. The most abundant RNA in this class would be 100 times more concentrated than the DNA sites to which it anneals. Thus, no more than 1% of this species could be bound and at this low concentration the final yield would undoubtedly be much lower. For this species the reaction would be first order at a rate determined by RNA concentration rather than DNA concentration Eq. (7). However, this RNA species would be about 1000 times less concentrated than are the repetitive DNA sequences; if k_1 were the same in the two cases, the reaction would be correspondingly slower but still the fastest of any species in this class. At the other extreme in the unique class, the concentration of the least abundant RNA would be about equal to that of its DNA site. In this case the reaction would be second order with a rate proportional to concentration of both reactants.

The most important species to consider in the intermediate class are the rRNA molecules. Assume the 18 S and 28 S species each represent 40% of the total RNA mass and each is homologous to 100 DNA sites or about 0.02% of the DNA [the mouse has 3×10^9 nucleotide pairs per genome (*15*); so if 28 S RNA has a molecular weight of 10^6, then 100 genes for 28 S RNA would require 0.023% of the DNA]. Since the 28 S RNA represents 40% of the total RNA, it would be about 2000 times more concentrated than its DNA sites at RNA/DNA inputs of one. The 0.05% that could anneal would do so as fast or faster than even the RNA for the satellite DNA since the rate would only be proportional to rRNA concentration. While the rRNA label would be an insignificant fraction of the hybrid at this input ratio, it would become significant at very low

RNA/DNA inputs; however, at these ratios the concentration of repetitive DNA sequences would be many times greater than the concentrations of rRNA. Thus, in this particular case the RNA from the satellite DNA would react the fastest of any RNA/DNA input.

D. The Reaction Is Limited for Thermodynamic and Kinetic Reasons

At any RNA/DNA input, almost all RNA species react with first-order kinetics approaching equilibrium at rates proportional to the concentration of the excess component. However, it should be emphasized that most of these very rare reactants, such as the unique species, are *never* fully bound at the usual RNA and DNA concentrations used (<100 μg). More information is needed to examine this assertion, but for the reaction of RNA from the lactose operon, *lac* RNA, to *lac* DNA, from the defective phage, ϕ80*dlac*, it was observed that the equilibrium yield of hybrid is half maximal when 1.5 μg RNA are allowed to react with 3 μg ϕ80 *dlac* DNA on a filter (Section II). This included 0.003 μg *lac* RNA and 0.17 μg *lac* DNA. In the preceding example, the most abundant RNA from the unique class of mouse DNA could have only 2×10^{-4} μg per 10 μg RNA input. It would be allowed to react with 2×10^{-6} μg of its DNA site per 10 μg DNA input. Thus, assuming that the association constants for *lac* RNA·DNA are comparable to unique mouse RNA·DNA, only a very small fraction of the DNA for this species would ever be filled. The same would hold for all the unique DNA sites in this range of reactant concentrations.

Thus, in most eucaryotic cell RNA·DNA hybridizations, the major species being followed are the repetitive DNA sequences; this follows from thermodynamic as well as kinetic reasoning. This would explain why the annealing reactions of mammalian cells often show *no* sign of increase after 16 hours (*47, 75, 107*). For example, in the re-annealing of *Drosophila* DNA, about 20% of the DNA fragments (5 μg) were bound to 50 μg of filterbound DNA in 16–20 hours, but this fraction did not increase in 72 hours. Only when the concentrations of the reactants were increased to 1 mg and 2.8 to 3.7 mg, respectively, did the fraction bound continue to increase for many days (*107*). It seems likely that this fraction would never have become bound at the lower concentrations. It would also explain why the hybridization of animal cell rRNA is extremely inefficient. An examination of saturation curves reveals that 6% binding of cell rRNA is rarely achieved even in the presence of excess DNA (*24, 73, 74, 76, 77*), although the reaction appears to be complete in 16 hours (*73, 74*). In contrast, it is not difficult to bind 65% of bacterial rRNA to

ɔs the concentration of repetitive greater than the concentrations ıe RNA from the satellite DNA input.

ıodynamic and

NA species react with first-order ɔs proportional to the concen-r, it should be emphasized that s the unique species, are *never* oncentrations used (<100 μg). this assertion, but for the re-ıc RNA, to *lac* DNA, from the ed that the equilibrium yield tNA are allowed to react with I). This included 0.003 μg *lac* ng example, the most abundant A could have only 2×10^{-4} μg d to react with 2×10^{-6} μg of , assuming that the association e to unique mouse RNA·DNA, ıis species would ever be filled. DNA sites in this range of re-

NA hybridizations, the major DNA sequences; this follows easoning. This would explain n cells often show *no* sign of example, in the re-annealing fragments (5 μg) were bound s, but this fraction did not in-trations of the reactants were tively, did the fraction bound seems likely that this fraction ıwer concentrations. It would imal cell rRNA is extremely ırves reveals that 6% binding presence of excess DNA (*24*, rs to be complete in 16 hours ind 65% of bacterial rRNA to

its DNA (*43*, *71*). In the former case, the concentrations of the ribosomal gene sites are several hundred times lower than in the bacterial case.

E. Presaturation Competition

There can be extensive mismatching when nucleic acids from eucaryotic cells are annealed. Most investigators wish to estimate degrees of similarity between RNA populations from different cells or tissues. This is usually done by competing one RNA against another for a single DNA. Naturally, the mismatching of both the unlabeled competitor and the labeled RNA will tend to complicate any interpretation of the results. One modification of the procedure that has been used entails hybridization of the unlabeled competitor RNA *before* the labeled RNA. This procedure has been termed "presaturation" competition (*108*, *109*). Soeiro and Darnell (*108*) showed that less labeled RNA is excluded by presaturation than by "simultaneous" competition with an unlabeled competitor RNA. However, McCarthy *et al.* (*45*, *92*) found no difference in the two procedures with respect to exclusion of labeled RNA. Presaturation competition is favored by Soeiro and Darnell because they believe that once formed, the hybrid complex is stable to dilution. In contrast, mismatched hybrids are bound imperfectly and reversibly.

In theory, this modification should be of value to distinguish perfect from imperfect hybrids if the concentration of hybrid is such that the degree of reversibility in the absence of excess RNA is a function of the stability of the hybrid. In Section II, the reversibility of *E. coli* RNA·DNA hybrids was shown. Even though such hybrids are very stable as judged by high melting temperature, a large fraction of RNA could become dissociated from the DNA upon reincubation of the hybrid in the absence of free RNA. The fraction of RNA dissociated was a function of two related variables: the dissociation constant K and the concentration of hybrid (Fig. 3); i.e., it obeyed the law of mass action. Therefore, it follows that those [^{1}H]RNA hybrids for which the equilibrium constant is higher will tend to dissociate more upon reincubation and be replaced by [^{3}H]RNA that might or might not give more stable complexes. An [^{3}H]RNA complex might be less stable than its competitor but displace the [^{1}H] complex if the [^{3}H]RNA concentration were sufficiently high.

If the [^{1}H]RNA that competes with the [^{3}H]RNA for a common DNA site is present in great excess over the [^{3}H] species, it will tend to exclude ^{3}H from the hybrid in a "simultaneous" hybridization even though it is more poorly matched. However, in a "pre-saturation" hybridization the same relatively unstable [^{1}H] hybrids would tend to exchange with the

[^{3}H] competitors during the second incubation. Thus, prehybridization of the competitor tends to magnify the competition by those [^{1}H]RNA species that are the most stable or perfectly matched, and reduce the importance of relative [^{1}H/^{3}H] abundance. [^{1}H]RNA hybrids that might compete when present at the same time as the [^{3}H]RNA could be lost. The extent of competition will also be a function of the amount of DNA reacting. This follows from the fact that all hybrid complexes are reversible to varying degrees and thus there is a concentration dependence for their persistence even when perfectly matched (Fig. 3).

In some cases the same degree of competition would be achieved with simultaneous competition simply by lowering the ratio of [^{1}H]RNA to [^{3}H]RNA. Finally, it should be clear that with so many variables contributing to the result, great caution should be exercised in assigning a certain degree of relatedness for such a procedure. It is not surprising that some authors, comparing "simultaneous" to "prehybridization" competition, find differences (*108*) while other authors do not (*45*, *92*). A more rigorous analysis of each situation would be helpful.

VIII. Conclusions

Nucleic acid hybrid formation has been used as a research tool for approximately one decade. It has been a decade of remarkable progress toward understanding mechanisms of gene expression. Probably no other single procedure has contributed more to this success than has hybridization. However, the procedure is still usually used as a crude assay for genetic relatedness; less often is it used for precise quantitative estimations. Its value as a technique will become even greater when more careful studies are undertaken to determine the variables that contribute to final yield of hybrid. An attempt has been made here to consider some of these in a very general way.

Hybridization will undoubtedly continue to be of great value for some time. The additional complexity of the genomes of higher organisms gives added complexity to interpretations of hybridizations with their nucleic acids. One technical problem will be to achieve sufficient concentrations of the unique DNA sequences in order to study their function. An indirect approach would be provided if small segments of eucaryotic chromosomes could be purified; this is now possible for bacterial genes.

These and other problems will have to be solved by more rigorous analyses and by refinements in the techniques. This will be necessary unless a chemical method more useful than hybridization is developed to study nucleic acid specificity. At this time, no such method appears imminent.

ncubation. Thus, prehybridization
ıe competition by those [^{1}H]RNA
erfectly matched, and reduce the
ınce. [^{1}H]RNA hybrids that might
ne as the [^{3}H]RNA could be lost.
a function of the amount of DNA
that all hybrid complexes are
; there is a concentration depen-
perfectly matched (Fig. 3).
mpetition would be achieved with
wering the ratio of [^{1}H]RNA to
hat with so many variables con-
ıould be exercised in assigning a
a procedure. It is not surprising
eous" to "prehybridization" com-
ther authors do not (*45*, *92*). A
ı would be helpful.

sions

ıeen used as a research tool for
ı decade of remarkable progress
ıe expression. Probably no other
:o this success than has hybrid-
usually used as a crude assay
used for precise quantitative
)ecome even greater when more
ne the variables that contribute
:en made here to consider some

inue to be of great value for
ıe genomes of higher organisms
s of hybridizations with their
l be to achieve sufficient con-
in order to study their function.
if small segments of eucaryotic
w possible for bacterial genes.
o be solved by more rigorous
ıiques. This will be necessary
an hybridization is developed
ime, no such method appears

Acknowledgments

Unpublished experiments were performed by Thom Schwartz. Dr. Alfred Holtzer, Department of Chemistry, made important suggestions for Section II. This work was supported by Research Grant GM 09830 and by a Public Health Service Research Career Development Award (1 K04 GM 06688) from the National Institute of Medical Sciences. Support for the Washington University computing facilities was provided by Grant G-22296 from the National Science Foundation.

References

1. C. L. Schildkraut, J. Marmur, J. Fresco, and P. Doty, *J. Biol. Chem.* **236**, PC3 (1961).
2. B. D. Hall and S. Spiegelman, *Proc. Nat. Acad. Sci. U.S.* **47**, 137 (1961).
3. J. Marmur, R. Rownd, and C. L. Schildkraut, this Series **1**, 231 (1963).
4. G. Felsenfeld and H. T. Miles, *Annu. Rev. Biochem.* **36**, Part II, 407 (1967).
5. D. M. Crothers and B. H. Zimm, *J. Mol. Biol.* **9**, 1 (1964).
6. P. O. P. Ts'o, G. K. Helmkamp, and C. Sander, *Biochim. Biophys. Acta.* **55**, 584 (1962).
7. J. Bonner, G. Kung, and I. Bekhor, *Biochemistry* **6**, 3650 (1967).
8. B. L. McConaughy, C. D. Laird, and B. J. McCarthy, *Biochemistry* **8**, 3289 (1969).
9. J. Eigner, H. Boedtker, and G. Michaels, *Biochim. Biophys. Acta* **51**, 165 (1961).
10. G. Attardi, P. C. Huang, and S. Kabat, *Proc. Nat. Acad. Sci. U.S.* **53**, 1490 (1965).
11. J. Marmur and P. Doty, *J. Mol. Biol.* **5**, 109 (1962).
12. J. A. Subirana and P. Doty, *Biopolymers* **4**, 171 (1966).
13. K. J. Thrower and A. R. Peacocke, *Biochim. Biophys. Acta* **119**, 652 (1966).
14. J. G. Wetmur and N. Davidson, *J. Mol. Biol.* **31**, 349 (1968).
15. R. J. Britten and D. E. Kohne, *Science* **161**, 529 (1968).
16. A. Rich, *Proc. Nat. Acad. Sci. U.S.* **46**, 1044 (1960).
17. M. N. Lipsett, L. A. Heppel, and D. F. Bradley, *J. Biol. Chem.* **236**, 857 (1961).
18. M. N. Lipsett, *J. Biol. Chem.* **239**, 1256 (1964).
19. A. M. Michelson and C. Monny, *Biochim. Biophys. Acta* **149**, 107 (1967).
19a. S. K. Niyogi and C. A. Thomas, Jr., *J. Biol. Chem.* **243**, 1220 (1969).
20. B. J. McCarthy, *Bacteriol. Rev.* **31**, 215 (1967).
21. S. K. Niyogi and C. A. Thomas, Jr., *Biochem. Biophys. Res. Commun.* **26**, 51 (1967).
21a. S. K. Niyogi, *J. Biol. Chem.* **244**, 1576 (1969).
22. P. M. B. Walker, this Series **9**, 301 (1969).
23. C. A. Thomas, Jr., this Series **5**, 315 (1966).
24. R. L. Moore and B. J. McCarthy, *Biochem. Genet.* **2**, 75 (1968).
25. E. K. F. Bautz and F. A. Bautz, *Proc. Nat. Acad. Sci. U.S.* **52**, 1476 (1964).
26. T. Kotaka and R. L. Baldwin, *J. Mol. Biol.* **9**, 323 (1964).
27. B. J. McCarthy and B. L. McConaughy, *Biochem. Genet.* **2**, 37 (1968).
27a. O. Uhlenbeck, R. Harrison, and P. Doty, *in* "Molecular Associations in Biology" (B. Pullman, ed.), p. 107. Academic Press, New York, 1968.
28. D. Gillespie, *in* "Methods in Enzymology," Vol. 11, Part B (L. Grossman and K. Moldave, eds.), p. 641. Academic Press, New York, 1968.
29. E. T. Bolton and B. J. McCarthy, *Proc. Nat. Acad. Sci. U.S.* **48**, 1390 (1962).

30. D. Gillespie and S. Spiegelman, *J. Mol. Biol.* **12**, 829 (1965).
31. D. Kennell and A. Kotoulas, *J. Mol. Biol.* **34**, 71 (1968).
32. G. Bernardi, *Nature* **206**, 779 (1965).
33. P. M. B. Walker and A. McLaren, *Nature* **208**, 1175 (1965).
34. Y. Miyazawa and C. A. Thomas, Jr., *J. Mol. Biol.* **11**, 223 (1965).
35. D. E. Kohne, *Biophys. J.* **8**, 1104 (1968).
35a. D. E. Graham, *Anal. Biochem.* in press (1970).
36. W. Szybalski, *Proc. 8th Can. Cancer Conf.* **8**, 183 (1969).
37. R. F. Baker and C. Yanofsky, *Proc. Nat. Acad. Sci. U.S.* **60**, 313 (1968).
38. F. Imamoto, *Proc. Nat. Acad. Sci. U.S.* **60**, 305 (1968).
39. G. Attardi, S. Naono, J. Rouviere, F. Gros, and F. Gros, *Cold Spring Harbor Symp. Quant. Biol.* **28**, 363 (1963).
40. M. Hayashi, S. Spiegelman, N. C. Franklin, and S. E. Luria, *Proc. Nat. Acad. Sci. U.S.* **49**, 729 (1963).
41. J. R. Beckwith, E. R. Singer, and W. Epstein, *Cold Spring Harbor Symp. Quant. Biol.* **31**, 393 (1966).
42. S. Gottesman and J. R. Beckwith, *J. Mol. Biol.* **44**, 117 (1969).
43. D. Kennell, *J. Mol. Biol.* **34**, 85 (1968).
44. B. J. McCarthy and E. T. Bolton, *J. Mol. Biol.* **8**, 184 (1964).
45. R. B. Church and B. J. McCarthy, *Biochem. Genet.* **2**, 55 (1968).
46. M. A. Martin and B. H. Hoyer, *Biochemistry* **5**, 2706 (1966).
47. P. M. B. Walker and A. McLaren, *J. Mol. Biol.* **12**, 394 (1965).
48. J. O. Bishop, *Biochem. J.* **116**, 223 (1970).
49. G. H. Pigott and J. E. M. Midgley, *Biochem. J.* **110**, 251 (1969).
50. T. Schwartz, E. Craig, and D. Kennell, *J. Mol. Biol.* in press (1970).
51. L. Leive and V. Kollin, *J. Mol. Biol.* **24**, 247 (1967).
52. E. Steers, Jr., G. R. Craven, and C. B. Anfinsen, *J. Biol. Chem.* **240**, 2478 (1965).
53. J. L. Brown, S. Kooragian, J. Katze, and I. Zabin, *J. Biol. Chem.* **241**, 2826 (1966).
54. T. Jones and E. P. Kennedy, *J. Biol. Chem.* **244**, 5981 (1969).
55. I. Zabin, *Cold Spring Harbor Symp. Quant. Biol.* **28**, 431 (1963).
56. J. H. Miller, K. Ippen, J. G. Scaife, and J. R. Beckwith, *J. Mol. Biol.* **38**, 413 (1968).
57. N. Ledinko, *J. Mol. Biol.* **9**, 834 (1964).
58. W. Szybalski, *in* "Methods in Enzymology," Vol. 12, Part B (L. Grossman and K. Moldave, eds.), p. 330. Academic Press, New York, 1968.
59. J. Weigle, M. Meselson, and K. Paigen, *J. Mol. Biol.* **1**, 379 (1959).
60. L. G. Caro, *Virology* **25**, 226 (1965).
61. R. Lavallé and G. De Hauwer, *J. Mol. Biol.* **37**, 269 (1968).
62. P. L. Gage, Ph.D. thesis, Department of Biophysics, University of Chicago, 1970.
63. J. M. Sturtevant, S. A. Rice, and E. P. Geiduschek, *Disc. Faraday Soc.* **25**, 138 (1958).
64. M. A. Rawitscher, P. D. Ross, and J. M. Sturtevant, *J. Amer. Chem. Soc.* **85**, 1915 (1963).
65. D. M. Crothers, *Accounts Chem. Res.* **2**, 225 (1969).
66. M. P. Printz and P. H. von Hippel, *Proc. Nat. Acad. Sci. U.S.* **53**, 363 (1965).
67. J. A. Subirana, *Biopolymers* **4**, 189 (1966).
68. A. P. Nygaard and B. D. Hall, *J. Mol. Biol.* **9**, 125 (1964).
69. J. O. Bishop, *Biochem. J.* **113**, 805 (1969).

ol. 12, 829 (1965).
34, 71 (1968).

ure 208, 1175 (1965).
Mol. Biol. 11, 223 (1965).
).
(1970).
f. 8, 183 (1969).
cad. Sci. U.S. 60, 313 (1968).
30, 305 (1968).
os, and F. Gros, *Cold Spring Harbor*

in, and S. E. Luria, *Proc. Nat. Acad.*

ein, *Cold Spring Harbor Symp. Quant.*

Biol. 44, 117 (1969).

Biol. 8, 184 (1964).
em. *Genet.* 2, 55 (1968).
try 5, 2706 (1966).
Biol. 12, 394 (1965).
).
hem. *J.* 110, 251 (1969).
Mol. Biol. in press (1970).
247 (1967).
Anfinsen, *J. Biol. Chem.* 240, 2478

I. Zabin, *J. Biol. Chem.* 241, 2826

. 244, 5981 (1969).
t. *Biol.* 28, 431 (1963).
. R. Beckwith, *J. Mol. Biol.* 38, 413

" Vol. 12, Part B (L. Grossman and
s, New York, 1968.
Mol. Biol. 1, 379 (1959).

37, 269 (1968).
physics, University of Chicago, 1970.
luschek, *Disc. Faraday Soc.* 25, 138

Sturtevant, *J. Amer. Chem. Soc.* 85,

(1969).
at. *Acad. Sci. U.S.* 53, 363 (1965).

. 9, 125 (1964).

70. J. O. Bishop, *Nature* 224, 600 (1969).
71. R. J. Avery and J. E. M. Midgley, *Biochem. J.* 115, 383 (1969).
72. M. Oishi and N. Sueoka, *Proc. Nat. Acad. Sci. U.S.* 54, 483 (1965).
73. F. M. Ritossa and S. Spiegelman, *Proc. Nat. Acad. Sci. U.S.* 53, 737 (1965).
74. H. Wallace and M. L. Birnstiel, *Biochim. Biophys. Acta* 114, 296 (1966).
75. J. R. Little and H. A. Donahue, *Proc. Nat. Acad. Sci. U.S.* in press (1970).
76. E. H. McConkey and J. W. Hopkins, *Proc. Nat. Acad. Sci. U.S.* 51, 1197 (1964).
77. K. Matsuda and A. Siegel, *Proc. Nat. Acad. Sci. U.S.* 58, 673 (1967).
78. J. O. Bishop, F. W. Robertson, J. A. Burns, and M. Melli, *Biochem. J.* 115, 361 (1969).
79. F. Imamoto, V. Morikawa, K. Sato, S. Misima, and T. Nishimura, *J. Mol. Biol.* 13, 157 (1965).
80. A. Landy, H. Abelson, M. Goodman, and J. D. Smith, *J. Mol. Biol.* 29, 457 (1967).
81. T. Ando and H. Ozeki, *Proc. Nat. Acad. Sci. U.S.* 59, 792 (1968).
82. D. E. Morse, R. Mosteller, R. F. Baker, and C. Yanofsky, *Nature* 223, 40 (1969).
83. R. J. Avery, J. E. M. Midgley, and G. H. Pigott, *Biochem. J.* 115, 40 (1969).
84. C. M. Guldberg and P. Waage, *in* "The Law of Mass Action," A centenary volume, Chapter 1. Norske Videnskaps-Akademi, Oslo, 1864–1964.
85. B. D. Hall, A. P. Nygaard, and M. H. Green, *J. Mol. Biol.* 9, 143 (1964).
86. A. Bolle, R. H. Epstein, W. Salser, and E. P. Geiduschek, *J. Mol. Biol.* 31, 325 (1968).
87. L. P. Gage, D. Wilson, and E. P. Geiduschek, *Fed. Proc.* 26, 873 (1967).
88. M. Denis, *J. Mol. Biol.* 22, 285 (1966).
89. B. J. McCarthy and B. H. Hoyer, *Proc. Nat. Acad. Sci. U.S.* 49, 729 (1964).
90. H. C. Birnboim, J. J. Pène, and J. E. Darnell, *Proc. Nat. Acad. Sci. U.S.* 58, 320 (1967).
91. V. R. Glišin, M. V. Glisin, and P. Doty, *Proc. Nat. Acad. Sci. U.S.* 56, 285 (1966).
92. A. H. Whiteley, B. J. McCarthy, and H. R. Whiteley, *Proc. Nat. Acad. Sci. U.S.* 55, 519 (1966).
93. J. E. Darnell, Jr., *Bacteriol. Rev.* 32, 262 (1969).
94. M. Waring and R. J. Britten, *Science* 154, 791 (1966).
95. F. Jacob and J. Monod, *J. Mol. Biol.* 3, 318 (1961).
96. R. J. Britten and E. H. Davidson, *Science* 165, 349 (1969).
97. R. Soeiro, H. C. Birnboim, and J. E. Darnell, *J. Mol. Biol.* 19, 362 (1966).
98. R. W. Shearer and B. J. McCarthy, *Biochemistry* 6, 283 (1967).
99. R. B. Church and B. J. McCarthy, *Proc. Nat. Acad. Sci. U.S.* 58, 1548 (1967).
100. G. Attardi, H. Parnas, M-l. H. Hwang, and B. Attardi, *J. Mol. Biol.* 20, 145 (1966).
101. J. Bonner and J. Widhohn, *Proc. Nat. Acad. Sci. U.S.* 57, 1379 (1967).
102. I. Bekhor, G. M. Kung, and J. Bonner, *J. Mol. Biol.* 39, 351 (1969).
103. R. C. C. Huang and P. C. Huang, *J. Mol. Biol.* 39, 365 (1969).
104. E. K. F. Bautz and F. A. Bautz, *Proc. Nat. Acad. Sci. U.S.* 52, 915 (1964).
105. H. Denis, *J. Mol. Biol.* 22, 269 (1966).
106. J. D. Watson, "Molecular Biology of the Gene." Benjamin, New York, 1965.
107. T. D. Entingh, *Genetics* in press (1970).
108. R. Soeiro and J. E. Darnell, *J. Mol. Biol.* 44, 551 (1969).
109. T. Kasai and E. K. F. Bautz, *in* "Organizational Biosynthesis" (H. J. Vogel, J. O. Lampen, and V. Bryson, eds.), p. 111. Academic Press, New York, 1967.